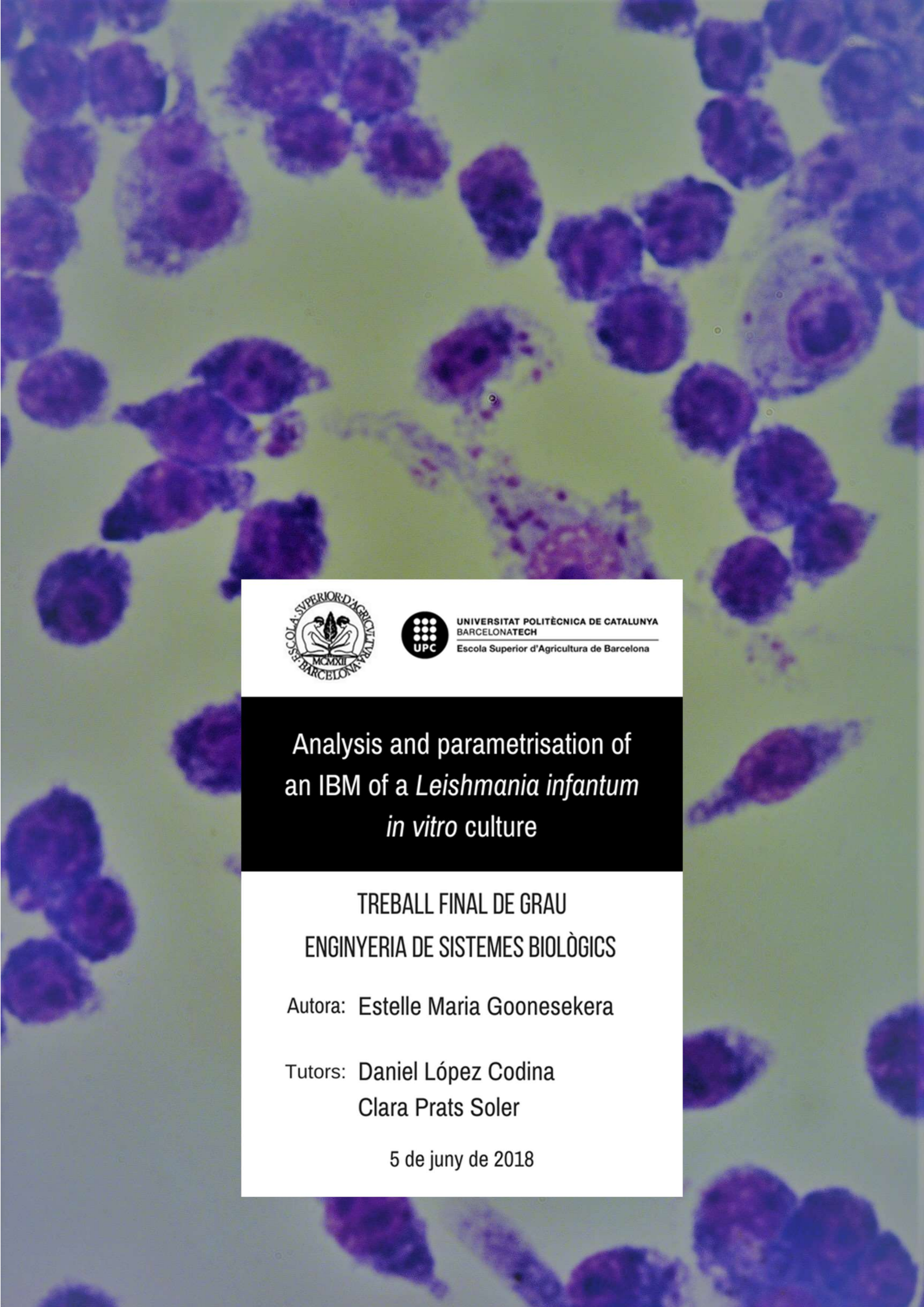


UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Escola Superior d'Agricultura de Barcelona

Analysis and parametrisation of an IBM of a *Leishmania infantum* *in vitro* culture

TREBALL FINAL DE GRAU
ENGINYERIA DE SISTEMES BIOLÒGICS

Autora: Estelle Maria Goonesekera

Tutors: Daniel López Codina
Clara Prats Soler

5 de juny de 2018

Abstract

Leishmaniasis is a vector borne protozoan parasitic disease with the ninth highest burden amongst infectious diseases. There is a pressing need for the development of novel, improved drugs for its treatment. Currently, the knowledge of the mechanisms of action of the parasite and the drugs that treat it is still limited. As a consequence, drug research begins with *in vitro* screening. This makes an increase in the understanding of *Leishmania* infected macrophage cultures necessary for the improvement of drug screening methods.

Aim:

The purpose of this work is to parametrise an individual based model (IBM) of a *Leishmania infantum* and murine cell line RAW 264.7 *in vitro* culture. This process consists of the comparison of experimental results and the simulation outcome of the IBM. It is a necessary process in order to adjust the parameters of the model so that it gives the closest representation of the real system. On the one hand, this serves to shed light on the behaviour of this particular culture. On the other hand, the work proposes a mathematical methodology for parametrisation that is extensible to other cultures (different strains, cell lines, parasite species and medium composition).

Experimental methods:

The experimental design used to obtain the experimental data for the parametrisation of the model is described. The percentage of infected macrophages and the number of parasites per cell was determined at several time-points after infection.

Parametrisation:

The mathematical methodology for parametrisation is described and applied to this particular model. A preliminary analysis of the model was carried out in order to determine, for each parameter, the interval of values where simulation outcome and experimental results were most similar. The combination of these intervals, the parameter space, was then sampled using the Latin Hypercube Sampling (LHS) technique. Finally, the sampled combinations were implemented in the model and the outcome analysed.

Results:

On the one hand, a mathematical methodology for the parametrisation process was developed, extensible to different cultures. One the other hand, the combination that best reproduces the experimental results was determined.

Conclusion:

The development of the IBM has the potential to increase our understanding of the system and, therefore, lead to a more accurate approach in understanding and evaluating drug assays. However, the model still requires further expansion. In spite of the satisfactory results, there are still some improvements to be made to the parametrisation methodology.

Resum

La leishmaniosi és una malaltia parasitària provocada per un protozou transmès per un vector. La seva incidència la situa com la novena malaltia infecciosa a nivell mundial. La necessitat de desenvolupar fàrmacs nous i millors és imperiosa. Actualment, encara, els coneixements sobre el mecanisme d'acció del paràsit i dels fàrmacs són limitats. Com a conseqüència, la recerca de nous fàrmacs s'inicia amb assajos *in vitro*. Això fa imprescindible per a la millora dels mètodes de triatge de fàrmacs una millor comprensió dels cultius de macròfags infectats per *Leishmania*.

Objectiu:

El propòsit d'aquest treball és parametritzar un model basat en l'individu (IBM) del cultiu *in vitro* de *Leishmania infantum* i la línia cel·lular murina RAW 264.7. Aquest procés implica la comparació de resultats experimentals amb les sortides del simulador i és necessari per a ajustar els paràmetres del model de manera que aquest representi amb màxima fidelitat el sistema real. Per una banda, serveix per a incrementar la comprensió del comportament d'aquest cultiu en particular. Per altra banda, el treball proposa una metodologia matemàtica per a la parametrització que pretén ser extensible a altres cultius (diferents soques, línies cel·lulars, espècies i composicions del medi).

Mètode experimental:

Es descriu el mètode experimental emprat per a obtenir les dades utilitzades per a la parametrització. Es determinen a diferents instants de temps el percentatge de macròfags infectats i el nombre de paràsits per cèl·lula.

Parametrització:

Es desenvolupa una metodologia matemàtica per a la parametrització i s'aplica a aquest model en particular. Un anàlisi preliminar permet establir en quins rangs de valors dels paràmetres la sortida del simulador s'assembla més a les dades experimentals. El conjunt d'aquests intervals, l'espai de paràmetres, és mostrejat amb la tècnica Latin Hypercube Sampling (LHS). Finalment, les combinacions de paràmetres mostrejades són implementades en el model i s'analitza el resultat.

Resultats:

Per una banda, s'ha desenvolupat una metodologia matemàtica per al procés de parametrització que és extensible a diferents cultius. Per altra banda, s'ha determinat la combinació de paràmetres mostrejats que millor reprodueix els resultats experimentals.

Conclusions:

El desenvolupament d'una metodologia per parametritzar models basats en l'individu té el potencial d'incrementar la nostra comprensió del sistema i, per tant, portarà a un enfocament més apropiat per a la comprensió i avaluació dels assajos de fàrmacs. Malgrat els bons resultats es constata que es poden realitzar, encara, algunes millores en la metodologia de parametrització.

Resumen

La Leishmaniasis es una enfermedad parasitaria provocada por un protozoo transmitido por un vector. Su incidencia la sitúa como la novena enfermedad infecciosa a nivel mundial. La necesidad de desarrollar fármacos nuevos y mejores es imperante. Actualmente, los conocimientos sobre el mecanismo de acción del parásito y los fármacos son aún limitados. Como consecuencia, la investigación de nuevos fármacos empieza con ensayos *in vitro*. Es necesaria para la mejora de los métodos de elección de fármacos una mejor comprensión de los cultivos de macrófagos infectados por *Leishmania*.

Objetivo:

El propósito de este trabajo es parametrizar un modelo basa en el individuo (IBM) del cultivo *in vitro* de *Leishmania infantum* y la línea celular murina RAW 264.7. Este proceso implica la comparación de resultados experimentales con las salidas del simulador y es necesario para ajustar los parámetros del modelo de manera que este represente con máxima fidelidad el sistema real. Por un lado, sirve para incrementar la compresión del comportamiento de este cultivo en particular. Por otro lado, el trabajo propone una metodología matemática para la parametrización que pretende ser extensible a otros cultivos (diferentes cepas, líneas celulares, especies y composiciones del medio).

Método experimental:

Se describe el método experimental utilizado para obtener los datos usados para la parametrización. Se determinan a diferentes instantes de tiempo el porcentaje de macrófagos infectados y el número de parásitos por célula.

Parametrización:

Se desarrolla una metodología matemática para la parametrización y se aplica a este modelo en particular. Un análisis preliminar permite establecer en qué rangos de los valores de los parámetros la salida del simulador es más parecida a los datos experimentales. El conjunto de estos intervalos, el espacio de parámetros, se muestrea con la técnica Latin Hypercube Sampling (LHS). Finalmente, las combinaciones de parámetros muestreados son implementados en el modelo y su resultado es analizado.

Resultados:

Por un lado, se ha desarrollado una metodología matemática para el proceso de parametrización que es extensible a diferentes cultivos. Por otro lado, se ha determinado la combinación de parámetros muestreados que mejor reproduce los resultados experimentales.

Conclusiones:

El desarrollo de un IBM tiene el potencial de incrementar nuestra comprensión del sistema y, por tanto, llevará a un enfoque más apropiado para la comprensión y evaluación de los ensayos de fármacos. Sin embargo, el modelo requiere expansión. A pesar de los buenos resultados, se constata que se pueden realizar mejoras en la metodología de parametrización.

Table of contents

Figures

Tables

Acknowledgements

I would first and foremost like to thank Daniel López and Clara Prats, my tutors, for their guidance, patience and for having confidence in me.

To Berta Raventós, much cited in this paper, without whose work this one would not even be possible. I express great admiration for the amount of time and effort spent on the development of model.

Next, I would like to thank Quim Valls for his immense help with the NetLogo platform. I would not have results if it were not for him.

To Maria Cristina Riera Lizandra, Roser Fisa and Maria Magdalena Alcover for their expertise in parasitary diseases.

To Martí Català for proposing, and repeatedly explaining, the LHS technique. To Nura Ahmad, for helping me actually carry out the technique. To Edneide Ramalho, for the hours she invested helping me with software that we never managed to make work.

This bachelor thesis has proved that one does not ever succeed alone. I am honoured to have been able to work with such talented researchers and such great people.

1. Introduction

1.1. Neglected tropical diseases

The term neglected tropical diseases (NTD) refers to seventeen diseases caused by parasites or bacteria that are prevalent in tropical and subtropical conditions (WHO 2017), as shown in Figure 1. The relationship between these illnesses is not biological, it lies in the fact that they have been mostly eradicated in the most developed parts of the world and persist only in the poorest and most marginalised communities and in areas of conflict (CDC 2011). NTDs affect more than a billion people in more than 149 territories. All low-income countries[1] are affected by at least five NTDs simultaneously. It is estimated that, every year, these diseases are responsible for 57 million DALYs[2] . It is also estimated that mass drug administration for most these diseases has a cost of less than 50 dollar cents per person and year (CDC 2011).

There exists a mutual dependence between neglected tropical diseases and poverty. The first would not be present on the scale that they are if the quality of life of those who suffer them was better and, in turn, the illnesses trap those who suffer them in the cycle of poverty. The diseases do not only affect people's short term health and wellbeing, they also influence physical and cognitive development, entail a social stigma that can lead to discrimination (especially in women and girls) and limit productivity, further worsening the state of poverty (CDC 2011).

Neglected tropical diseases are responsible for massive suffering and can lead to permanent disability and even death. Despite this, as the term *neglected* implies, they have scarce visibility and are far from being prioritised due to them almost exclusively affecting vulnerable communities without political representation. These communities generally live in rural areas or in suburbs in countries of the global South (WHO 2017).

Those communities most affected by NTDs are those that live in close proximity to vectors and domestic animals (WHO 2017). Many NTDs are vector borne diseases. In epidemiology, a disease vector is any agent (living or inert) that transmits a pathogen to an organism. The role of animal reservoirs is also important. A reservoir is any agent or substance that harbours an infectious agent while suffering neither its symptoms nor effects. Reservoirs act as a source of infection for other organisms. Many domestic animals act as reservoirs for NTDs and other vector borne diseases.

The World Health Organisation (WHO) proposes various strategies to combat these diseases: large-scale preventive treatments, more intensive management of cases in primary attention centres and vector control. It also includes general prevention methods such as access to safe drinking water and adequate sanitation and hygiene (WHO 2017).

[1] According to the Word Bank, a country is considered of low income if the national income per capita is inferior to 1005 $ per year.

[2] DALYs or disability adjusted life years are a measure of overall disease burden, expressed as the number of years lost due to ill-health, disability or early death. It was developed in the 1990s as a way of comparing the overall health and life expectancy of different countries.

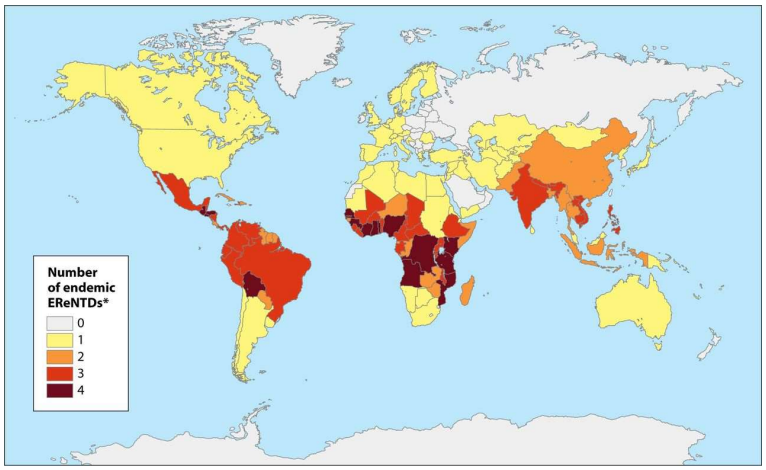


Figure 1. Global distribution of neglected tropical diseases (Mackey et al. 2014).

1.2.Leishmaniasis

Leishmaniasis, a notorious member of the NTDs, is a vector borne disease that may be caused by twenty species of the *Leishmania* genus, members of the protozoan Trypanosoma family. The parasites develop part of their cycle in humans and part in sand flies, which persists in warm and humid climates. This includes thirty species of mostly nocturnal phlebotomids with *Phlebotomus* being the most common in the Old World (Europe, Asia and Africa) and *Lutzomyia* in the New World (the Americas). The cycle of the parasite is shown in Figure 2. The parasite infects macrophagic cells of the mononuclear phagocytic system (skin, bone marrow, liver, spleen and lymphatic nodes).

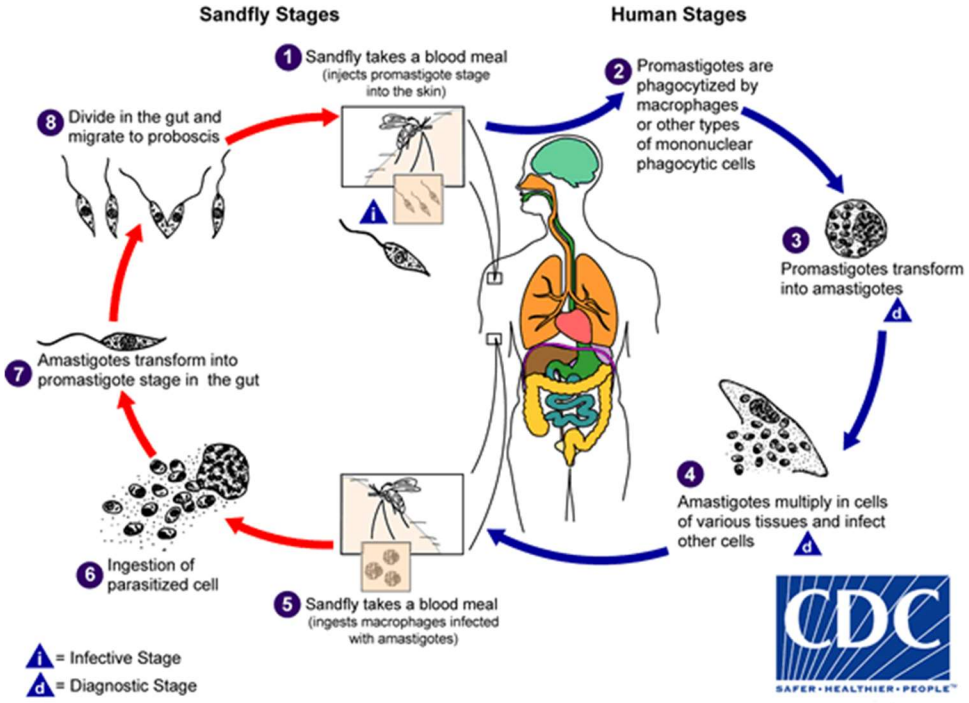


Figure 2. Lifecycle of Leishmania *(CDC 2011).*

The parasite has two distinct forms with physiological, biochemical and molecular differences. The extracellular form, known as promastigote (Figure 3A), is the one found in the digestive tract of the vector. It is mobile and has a flagellum. The intracellular form, known as amastigote (Figure 3B), is the one found in infected phagocytes. It is not mobile, has no flagellum and is much more virulent.

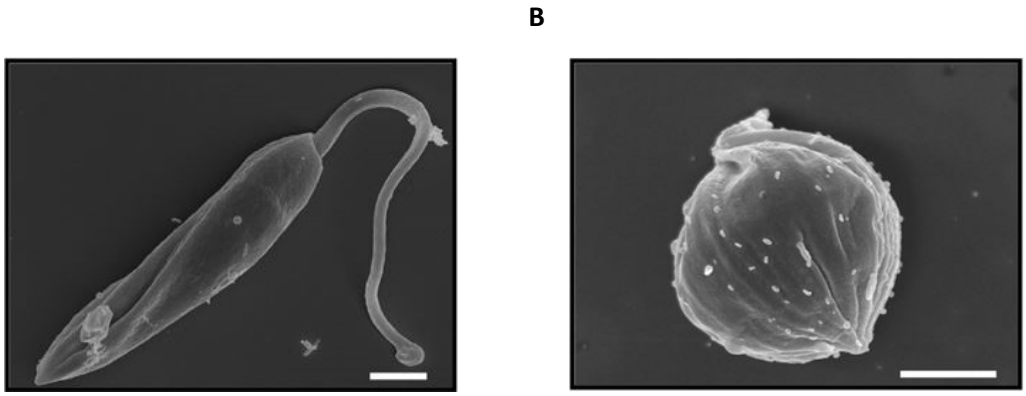


Figure 3. (A) Promastigote (extracellular) and (B) amastigote (intracellular) forms of Leishmania *(Samant et al. 2018).*

There are three forms of the illness, depending on the symptoms: cutaneous, mucocutaneous and visceral. The cutaneous form entails skin ulcers, the mucocutaneous form entails ulcers at the nose, mouth and other mucosal tissues and visceral Leishmaniasis not only entails ulcers, but fever, a decrease in red blood cells and inflammation of the spleen and liver. The cutaneous form is the most common and is not fatal but can lead to severe scarring. The visceral form is the most severe and can lead to death if left untreated. Its progression depends on the efficacy of the immune system and the virulence of the parasite. It may even become chronic.

1.2.1. Epidemiology

It is estimated that 4 – 12 million people in 98 countries are affected by Leishmaniasis, 200 million people live where the illness is common and it causes 20 – 50 thousand deaths a year (WHO 2017). Belonging to the group of neglected tropical diseases, its risk factors include: poverty, malnutrition, deforestation, lack of waste management and sanitation, lack of urbanisation and a depressed immune system. Leishmaniasis accounts for 2357000 DALYs (Hotez et al. 2004).

Control measures to avoid the spread of Leishmaniasis include investment in case finding and treatment, vector control with residual insecticide house spraying, insecticide impregnated nets for beds and animal reservoir control (WHO 2017). For now, sand flies are still very sensitive to insecticides (Desjeux 1996), but there is a risk of them developing resistances in the future.

1.2.2. Leishmaniasis and poverty

Leishmaniasis is generally considered to be one of the most neglected diseases. There is a link between poverty of sufferers and the consequent failure of a free market pharmaceuticals to address them (Yamey & Torreele 2002).

Poverty is associated with risk increasing factors such has poor housing, lack of sanitation and deficient waste management, coupled with a significant number of untreated patients and domestic animals. Poverty also increases the progression, morbidity and mortality of the disease. This is due to poor nutrition in combination with poor access to health care and the high

cost the process of prevention and treatment, from mosquito nets, to transport and the treatment itself (Alvar et al. 2006).

The issue can also be looked at from a gender perspective. Women suffer more stigma than men and, therefore, suffer more barriers when it comes to healthcare. For example, a woman may need a male relative to accompany her to receive treatment. Less access to healthcare and treatment mean that women bare a higher disease burden than men.

Poverty and Leishmaniasis create a mutually perpetuating cycle: poverty increases the likelihood of contracting the disease and makes it take a higher toll on the system. At the same time, the high cost of treatment and the loss of household income due to the inability to work or having to accompany a relative have long term consequences on the family economy. Basically, the disease increases poverty (Alvar et al. 2006).

Like all NTDs, Leishmaniasis is largely ignored due to the fact that the people that suffer it have neither voice nor power to influence decision making. This is especially significant because public investment is urgently needed to break the vicious cycle of disease and poverty. On the one hand, to improve the general quality of life of people and, on the other hand, to promote active detection and early treatment, as well as other control measures (Desjeux 1996).

1.3.Treatment of Leishmaniasis

1.3.1. Current status

There have been drugs to treat leishmaniosis since the 1960s and the most commonly used are pentavalent antimonials, amphotericin B and miltefosine. However, these are highly inadequate due to moderate to high toxicity, prohibitive costs and emergence of resistance (Berman et al. 1989). All are administered parenterally except miltefosine, which is not adequate for children nor pregnant women. Leishmaniasis has been called a major health problem with no satisfactory treatment so far (Mishra et al. 2007). There is a clear need for safe, potent, oral, affordable and effective drugs to treat the different forms of the illness.

1.3.2. Drug discovery

The effort in drug discovery for human Leishmaniasis is limited due to the poor economic returns related to the fact that the disease affects mostly impoverished endemic areas (Nagle et al. 2014). Organisations leading the research into drugs for Leishmaniasis treatment are the Drugs for Neglected Diseases initiative (DNDi), Genomics Institute of the Novartis Research Foundation (GNF) and the University of Dundee.

Currently, there are few lead compounds, possibly due to technological hurdles and/or lack of resources. A lead compound is a chemical entity that has pharmacological or biological activity and could potentially be developed into a drug. Nevertheless, its structure needs to be optimised in order to better fit its target, thereby maximizing its beneficial effects and minimizing side effects.

The compounds to be screened as possible leads against different diseases are stored in libraries. These are large collections of functionally diverse molecules that are used for the discovery of chemical matter active against a wide variety of targets and for phenotypic screening. The main challenges regarding the identification of novel leads are the difficulty of accessing high quality hits from screening libraries and a lack of a standardised screening cascade for the various stages of discovery and development (Bhuniya et al. 2015). All things considered, high throughput phenotypic screening essays offer a powerful tool for the discovery of therapeutically relevant leads for drug discovery (Moffat et al. 2017).

Drug discovery efforts can be based on targets or on phenotypic assays. The first are hypothesis-driven molecular approaches where mechanistic details necessary for their implementation are not always available. The latter consist of evaluating the drug against the observable characteristics of the organism, for example, its ability to survive.

Early in the progression of drug discovery, knowledge is acquired by empirical analysis. It serves to identify substances with therapeutic potential and their mechanism of action. However, even if this approach does not require understanding of the mechanism of action, it does require understanding of the biology enough to assess the effect of the therapeutic substance. Some notion of the dose-response relationship is necessary to further develop the substance as a drug (Swinney 2013).

Phenotypic assays, previously non-existent due to complexity of cultures and lack of resources, have relatively recently been made possible by advances in technology. For diseases where many of the molecular mechanisms are unknown, phenotypic assays have been more successful than target-based efforts in identifying potential drugs. They have also been useful for understanding the biology of largely unknown illnesses, identifying possible targets and agents acting on previously unknown targets (Swinney 2013; Nagle et al. 2014).

Once a lead has been identified, subsequent assays are performed in order to determine its suitability as a therapeutic agent. These include a mammalian cytotoxicity test, in order to discern between inhibitory activity of interest and general cytotoxicity. This is followed by a microsomal stability test to determine the effect of oxidative metabolism on the compound. Finally, an *in vivo* efficacy test is run in a murine model. It is important to screen appropriately in order to avoid non-specific effects.

In the case of Leishmaniasis, due to the various disadvantages of parenteral administration, there is a pressing need for oral drugs that can be administered to the general population (miltefosine is inadequate for pregnant women and children). Therefore, oral bioavailability is a trait being pursued in lead development (Ortiz et al. 2017).

Figure 4 shows the timeline of the drug development process, from the initial screening process until the final product is obtained. Notice how from thousands of compounds screened in the early stages, only one or two (approximately) are deemed suitable for administration. Note also that the period covers 11 to 15 years. This highlights the need to speed up initial screening in order to shorten the time needed to develop a drug, as well as to optimise resources.

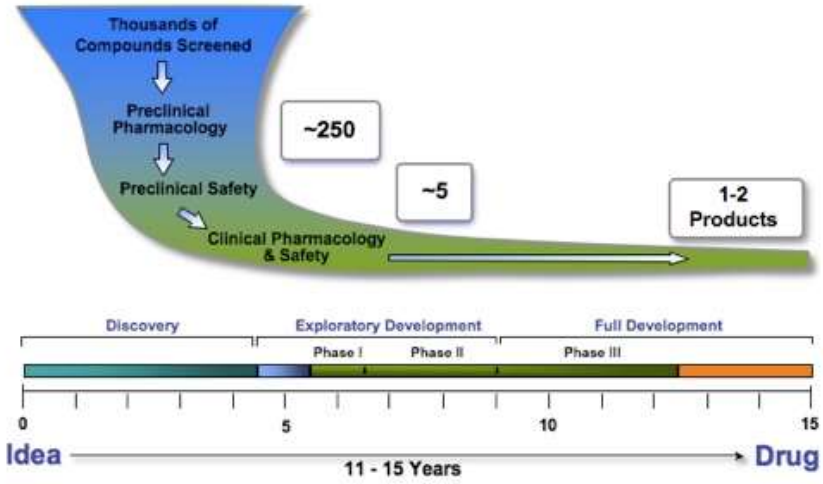


Figure 4. Drug development timeline, from initial screening to the obtaining of the final product.

Figure 5 shows a timeline of antileishmanial drug discovery, clearly showing its slow pace (Samant et al. 2018). Some of the most commonly used drugs (pentavalent antimonials and amphotericin B) have been used since the mid twentieth century. The gap between amphotericin B and miltefosine is more than forty years. It is also important to note that all of these drugs have major side effects and are far from ideal.

Figure 5. Timeline of antileishmanial drug discoveries from the early 1920s to 2017 (Samant et al. 2018).

The current trend in drug discovery for Leishmaniasis is drug repurposing, i.e., testing the leishmanicidal activity of drugs used to treat other diseases. For example, miltefosine is traditionally used as an anti-cancer drug. This approach offers a shorter and faster path towards drug identification and lowers the cost of development.

1.3.3. Phenotypic assays

In vitro cultures are commonly used for the preliminary screening of potential drugs. This technique requires less volume of substances, has a higher throughput and speed, lower cost and allows the reduction of the number of animals used in subsequent *in vivo* essays. However, it inevitably ignores important host factors such as absorption, immune response, etc. Ideally, *in vitro* cultures should contain an active (dividing) population of the mammalian stage of the parasite, have an easily quantified measure of drug activity and accurately show the activity of standard drugs at the approximate concentration achieved in tissues (Croft 1986).

High throughput screening (HTS) assays are implemented using an *in vitro Leishmania* culture. Usually, a single concentration of each substance is tested. It is validated by including a test for standard drugs. It has the potential of screening many compounds at once: many cultures are incubated in parallel, each in the presence of a potential lead compound. If correctly validated, can account for a high level of automation (Siqueira-Neto et al. 2010).

The techniques used for the quantification of drug activity are very diverse. There are three forms of culture to be used in this type of essay: promastigotes, axenic amastigotes and

intracellular amastigotes (Figure 6A, 6B and 6C, respectively). Because of the importance of culturing the appropriate parasite form under relevant physiological conditions (Nagle et al. 2014), there is an ongoing debate on the validity of especially promastigote and axenic amastigote essays, which will be subsequently described in detail. Differences in results between the three cultures have been widely reported.

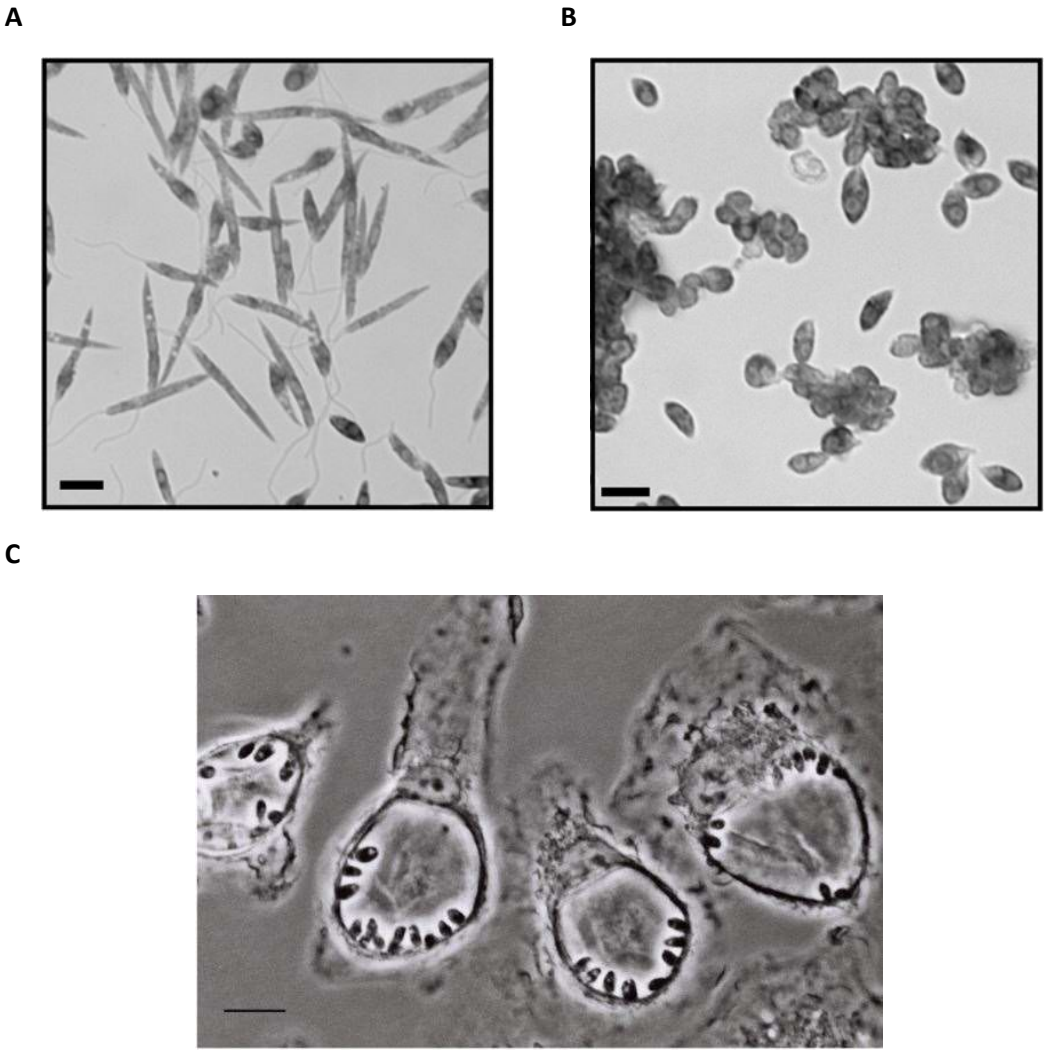


Figure 6. (A) Promastigote assay, (B) axenic amastigote assay and (C) intracellular assay (Yau et al. 2010; Aulner et al. 2013).

The promastigote culture consists of screening the drug against free promastigotes, without any mammalian cells, i.e., without infection. They are mostly used because they are easy to culture and of low cost, allowing a high screening throughput (Sharlow et al. 2009). It was observed by Carrió et al. (2000) that culture conditions affected the value of IC_{50}[3]. The main issue related to this type of culture is the high discrepancy found between its results and those of intracellular and *in vivo* assays (Callahan et al. 1997; Ephros et al. 1997), which is said to be related to the fact that the promastigote form is the one found in the insect vector and is therefore not relevant to the disease in mammals.

[3] Concentration of an inhibitory substance capable of producing 50% of the maximum inhibition possible *in vitro*.

Axenic amastigotes are more similar to the disease relevant parasite stage (Sereno et al. 2000), but they are cultured alone, not inside cells as they are found in mammals. This entails certain differences in protein expression and drug susceptibility (Nagle et al. 2014) with a consequent high false positive rate (De Rycker et al. 2013). It is used because it too permits a high screening throughput.

The intracellular amastigote assay is the most physiologically relevant, as it not only detects drug susceptibility but also host-cell dependent mechanisms. It is the most suitable of the three, even though it offers the lowest throughput due to its higher complexity and cost (Siqueira-Neto et al. 2010; Nagle et al. 2014; Nühs et al. 2015). It allows the elimination of compounds that do not affect amastigotes and also identifies those not effective on promastigotes that are effective on amastigotes.

Sensitivity of *Leishmania* to drugs depends on species (Nelson et al. 1979; Avila & Casanova 1982; Rangel et al. 1996; Carrió et al. 2000). This difference has been proposed to be due to differences in biochemical composition related to membrane sterol and lipid content (Goad et al. 1984; Beach et al. 1988; Beach et al. 1979).

Sensitivity also depends on lifecycle stage, in many cases being higher in promastigotes than in intracellular amastigotes. This could be due to stage-specific activity reflected in differences in physiology, such as differences in biochemical composition of the membrane (Bouazizi-Ben Messaoud et al. 2017). It could also be due to differences in division rate and metabolism or due to more or less exposure, depending on their extracellular or intracellular state (Moreira et al. 1998; Carrió et al. 2000; Escobar et al. 2002). The preparation of the culture and the compound have also been found to affect susceptibility (Ephros et al. 1997; Roberts et al. 1998). Berman et al. (1989) found that promastigotes and amastigotes have similar susceptibilities in the same environment, but it is clear that they do not exist in similar environments.

Even though in many cases the effect of drugs is higher on promastigotes than on amastigotes, there are certain cases when the inverse situation is found. This has been attributed to active concentration of drugs in the macrophage phagolysosome or parasite vacuole. This phenomenon has been observed, for instance, for antimony derivatives. It has also been attributed to metabolization of the drug inside the macrophage to a form with higher leishmanicidal activity and to intrinsic susceptibility of the intracellular stage to the drug in particular (Berman et al. 1987; Callahan et al. 1997; Ephros et al. 1997; Roberts et al. 1998). It has also been found that medium composition may modify promastigote susceptibility (Peixoto & Beverley 1987).

1.4.Individual based models

The aim of mathematical modelling is to find, interpret and validate approximate representations of systems. These are defined by concepts, elements and interactions that are described by means of mathematical objects and operations (Gomes Neves & Duarte Teodoro 2010). The main requirement of a model is that its results have to be consistent with experimental data (Feynman & Bosch 1967). Models should be coherent and have descriptive and predictive capacity. They are always simplifications, but they capture the essence of the system enough to address specific questions (Grimm & Railsback 2005).

There are two main strategies when it comes to modelling: top down and bottom up. If we are dealing with the modelling of a microbial community, the first describes the system at a population level and infers the behaviour of the individual from the macroscopic patterns. In the second, the behaviour of the system is described by establishing rules for the behaviour of the individuals and their interactions. Following these strategies, models can be classified into population and individual based models, respectively. The first, usually formalized by means of continuous equations, are more tractable, interpretable and understandable, but they can handle only a limited amount of complexity. Individual-Based Models (IBMs) can handle almost an unlimited amount of complexity, but are more difficult to develop, understand and explain. Neither type of model is entirely superior, all approaches are needed to maximize understanding. As is said in Peck (2004), “the world is complex and we need all the tools we can muster to understand it”.

Individual based models are those simulation models that treat individual as discrete and unique entities that have at least one property, apart from age, that changes during their life cycle (Grimm 1999). The behaviour of the individuals is governed by a set of rules (equations) and the overall behaviour of the system is a consequence of the interaction between individuals. This permits them to explore the relationship between the low level interactions of individuals and the macroscopic patterns that emerge from them (Ginovart et al. 2011).

In fact, an interesting aspect if IBMs are emergent behaviours at the population level, which are not evident nor self-contained, but that are observed none the less as a consequence of the behavioural patterns of the individuals (Grimm 1999). This kind of model aims to trace back the system’s properties to the behaviour of the individuals, explaining the system’s properties as a consequence of the properties of the single individuals (Kaiser 1979).

IBMs incorporate stochasticity at an individual level, in order to model individual variability. However, the macroscopic behaviour is deterministic: predictions can be made. There is an exception to this behaviour when working with few initial individuals, where, due to deterministic chaos, the macroscopic pattern is unpredictable. Together with the bottom up strategy, individual variability gives more realism to the model (Ferrer et al. 2008). The structure of IBMs also allows the integration of observations directly into the model without having to rewrite a whole set of equations (Grimm 1999).

This type of model is not easily constructed and is difficult to understand and analyse. Neither is the relevance of input information always clear. The necessary data at individual level is often difficult to obtain from experimental data, although individual observation techniques have improved in recent years. With these new technologies, traits and activities at an individual level can be observed and measured, for example, individually differing phenotypes in a microbial population. By combining this data with IBMs, the emergent population behaviour can be modelled.

Because the available data for the characterisation of individuals is not always available, a parametrisation process is always necessary. It consists of adjusting the parameters of a model in a way that the simulation outcome replicates the experimental results as closely as possible. This process often involves large numbers of parameters, making it very complex. There is often a range of values of a parameter where the simulation outcome is similar to that of the experimental system. Together, these intervals conform the parameter space, all the possible combinations of parameter values for a particular model.

Nowadays, the ODD protocol (Overview, Design concepts and Details) sets the standard for the description of IBMs and helps overcome their complexity, making their communication plainer. This helps to overcome the historical drawback of not having the necessary protocols for communication, resulting in a significant lack of clarity (Railsback 2001; Ferrer et al. 2008; Ginovart et al. 2011)

There are various situations in which IBMs are particularly useful. They allow the creation of models with explicit space and can integrate spatial heterogeneity (Ferrer et al. 2008). They are useful to study the impact of different elements involved in the model and to verify the validity of proposed mechanisms, contributing to the interpretation of data (Grimm 1999, Railsback 2001, Ferrer et al. 2008). Another interesting use is "simulation as experiment" (Peck 2004), a change in paradigm that could complement experimentation in those cultures that are costly, difficult or even impossible (Sharma 2005). All in all, IBMs are a good alternative to study complex systems (Grimm 1999), when these are formed by autonomous entities (Ferrer et al. 2008).

1.4.1. NetLogo

The platform used in this project for the development of the model is NetLogo (http://ccl.northwestern.edu/netlogo/), created in 1999 by Uri Wilensky. It is an open source software that is in continuous development at the Center for Connected Learning and Computer-Based Modeling of the United States (http://ccl.northwestern.edu/). It is a modelling environment specifically designed for the implementation of IBMs and the simulation of complex systems with explicit time and space. As a multiagent platform, it is capable of handling multiple independent, mobile agents.

The programming language is simpler than others used for the same purpose and it is much more visual and intuitive thanks to automatic animations and graphic controls. However, the rate of the simulation is slow compared to other platforms and languages.

1.4.2. Parasite models

The existing models related to Leishmaniasis are mostly epidemiological, modelling the transmission of the disease. Their main aim is to define control strategies. The main limitation, as with all levels of study in the field of NTDs, is the difficulty of parametrisation due to the lack of quality of the available data. To date, no *in silico* models describing the infection process have been published. However, there exist infection models of other parasitary illnesses, such as malaria and Chagas' disease.

There exists an IBM of the infection dynamics of *Plasmodium falciparum* (the parasite that causes malaria) in *in vitro* erythrocyte cultures developed by Ferrer et al. (2010). This model was used in order to define local interactions and integrate the complexity and stochasticity necessary to obtain the appropriate emergent behaviours. The model includes interactions between agents (erythrocytes and *P. falciparum*) and between the agents and their environment, including transport processes and the culture system.

Another model describing processes related to parasitary diseases is a cellular automaton of the development of Chagas' disease after stem cell transplantation, developed by Galvão et al. (2009). The model includes the different types of cells in cardiac tissue affected by Chagas' disease. Its aim is to comprehend the contribution of these different cells in chronic chagasic cardiomyopathy regeneration by validating behaviour hypotheses. These hypotheses are the rules that simulate cellular kinetics in the model.

1.5. Context and aim

1.5.1. Trajectory

This line of research started in 2015 with Núria Pedrós' bachelor thesis (Pedrós Barnils 2015): *Analysis and modelling of Trypanosomatidae family* in vitro *and* ex vivo *cultures*. Here, a continuous mathematical model for *in vitro* cultures was developed, characterising the growth of the culture as bi-lineal.

The development continued in 2016 with Berta Raventós' bachelor thesis (Raventós Roca 2016), where the aim was to improve the quantification methods of the experimental results from amastigote *in vitro* cultures. Different experimental methods were evaluated to assess their effects and improve the used methodology. A protocol for the digitalisation of the culture preparations was defined. As a result of the collection of images produced, a method for the automatization of the interpretation of images has recently been developed: Leishmaniasis Parasite Segmentation and Classification using Deep Learning (Górriz et al. 2018). This paper will be presented in the X Conference on Articulated Motion and Deformable Objects in Palma, Mallorca (Spain) on the 12-13 July 2018.

In 2017, Berta Raventós developed, in her master thesis, the IBM for *in vitro* cultures of *Leishmania infantum* that is further studied in this bachelor thesis (Raventós Roca 2017). The aim of the model is to describe the behaviour of such cultures and further the understanding of them by defining rules for the behaviour of macrophages and parasites. The experimental method was also defined in order to obtain results to validate the proposed mechanisms. A first approximation to parametrisation was carried out, which this bachelor thesis aims to continue.

1.5.2. The challenge: parameterising models of complex systems

Cultures where two or more agents (parasite species, chemical compounds, host cells…) interact are complex systems, where it is nearly impossible to evaluate the effect of each element on the behaviour of the system. Each agent behaves and interacts in its own way and the behaviour of the culture differs with strain and species of parasite, host cell line and medium composition, especially when a drug is added.

As shown in Figure 7, on the one hand, experiments are designed and their results collected and analysed. On the other hand, the model is developed based on proposed behaviour mechanisms and the parameters are set according to bibliographical data. In order to make the two approaches compatible, the experimental design must permit the measure of data that is useful for the comparison between experimental results and simulation output.

Those parameters whose values cannot be found in bibliography must be adjusted. This process is known as parametrisation and it involves the comparison of experimental data and simulation output with the aim of finding the parameter values that best replicate the experimental system. The steps of parametrisation are highlighted in orange in Figure 7. With the results produced, the final steps will be carried out and a global analysis will be possible in order to find the

combination which best fits the real system. This information will allow, in turn, the optimisation of both the experimental design and the IBM.

1.5.3. Aim and outline

The aim of this bachelor thesis is to develop a mathematical methodology for the parametrisation of the mathematical model describing the behaviour of a *Leishmania infantum* and cell line RAW 264.7 *in vitro* culture. With this methodology, the model is evaluated and the parameter space determined is explored, as highlighted in Figure 7. At the same time, it also strives to further understand the characteristic parameters of this particular culture, specifically those parameters that are difficult or impossible to measure experimentally.

A major challenge for the development of a parametrisation method that yields optimum results is the definition of strategies for the comparison of experimental and simulated systems. Considering the complexity of the real system and the large number of parameters involved in describing it, this challenge is by no means trivial. The developed methodology aims to overcome this hurdle and be extensible to different strains, species of parasite and host cell lines.

In Section 2, the experimental protocol for the obtainment of the experimental data used for the parametrisation is presented. Next, the model itself is described, followed by the description of its parametrisation process. The initial evaluation of the model results in the determining of the parameter space, which is subsequently explored. Finally, the results of both these stages are presented and discussed.

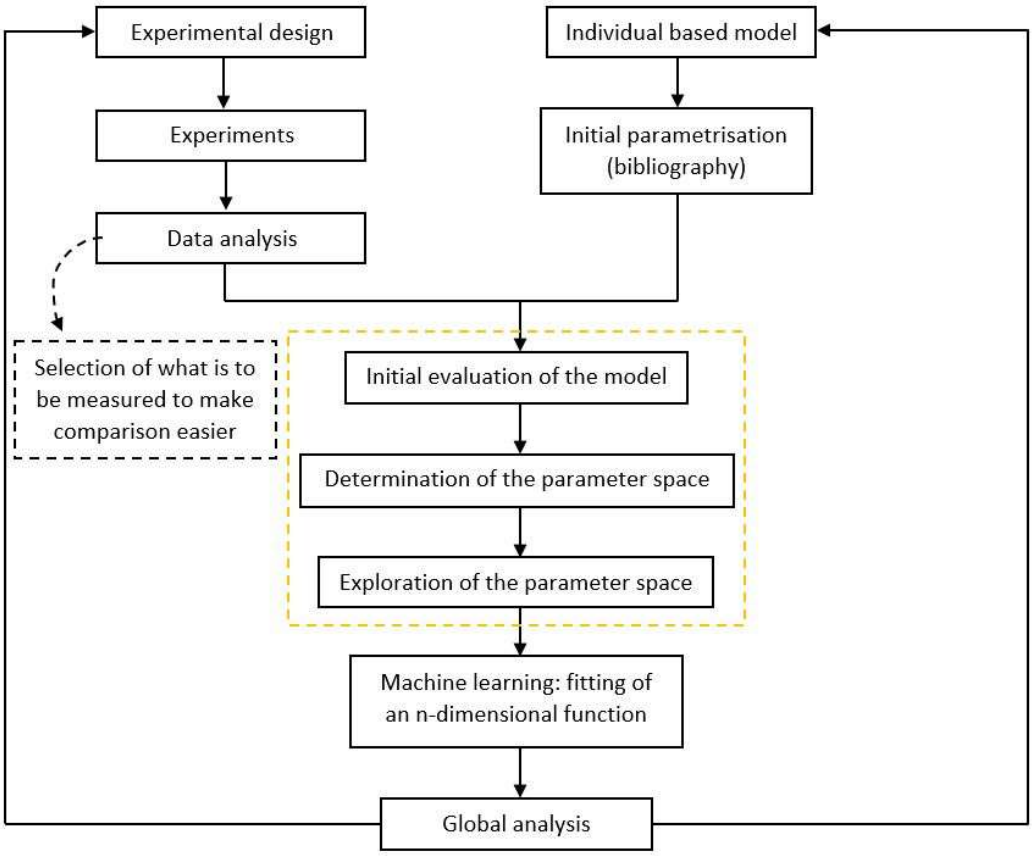


Figure 7. Flowchart of the steps necessary for the development of an IBM and its corresponding experimental design. In orange, the steps carried out in this bachelor thesis.

2. Materials and methods

The following experimental methods were developed and carried out by Raventós Roca (2017). All procedures that involve manipulation of cell lines and *Leishmania* took place in laboratories of Security Level II.

2.1.*Leishmania* promastigote culture

The strain used for the assay was *L. infantum* (WHOM/ES/2016/BCN-877) from the Cryobank of *Leishmania* strains pertaining to the Laboratory of Parasitology of the Facultat de Farmàcia i Ciències de l'Alimentació (University of Barcelona).

The promastigotes were cultivated in Schneider medium (Sigma, ref.: S9895-1L) supplemented with 20 % Foetal Bovine Serum (FBS) (AntibodyBcn), 1 % of Gentamicin (Sigma, ref.: G1397-10ML) and 1 % filtered human urine. The culture was kept in flasks of 25 cm^3 (Greiner bio-one; ref.: 658175) and maintained at a temperature of 26 ºC.

2.2.Cell line RAW 264.7 culture

The cell line used for the assay was the continuous cell line RAW 264.7. It was originally isolated from an induced tumour in a mouse by the intraperitoneal[4] injection of Abelson's Murine Leukaemia Virus (A-MuLV). The macrophages of this cell line are semi-adherent and find themselves in a permanent division process.

The culture was initiated from a vial of the original culture (10^6 cells/mL) frozen at - 80 °C in RPMI 1640 medium (Sigma, ref.: R7388-500ML) and 10% dimethyl sulfoxide (DMSO). In a 25 cm^3 flask, 1 mL of cell culture was deposited, together with 10 mL of RPMI 1640 medium supplemented with 10 % FBS and 1 % Penicillin/Streptomycin (P/S) (Sigma, ref.: P4333-20ML).

The culture was incubated at 37 °C and 5 % CO_2 so that the cells would divide forming a monolayer. The maintenance of the culture requires the removal of the contents of the flask, the adding of fresh medium and the energic scraping of the flasks surface with a Cell Scraper (Greiner bio-one, ref.: 541070) in order to separate the cells. This procedure is repeated every 3 or 4 days.

2.3.*In vitro* infection of the cell line with promastigotes

The *in vitro* infection of the cell line RAW 264.7 with *L. infantum* promastigotes took place in chambers slides (Thermo Scientific™ Nunc™ Lab-Tek™ II Chamber Slide™). This system is made up of eight removable polystyrene chambers attached to a glass microscope slide treated to ensure a uniform culture surface. The slide (25 x 27 mm) also allows the staining and microscopic examination of the culture. Five intervals were studied: 12 h, 24 h, 36 h, 48 h and 72 h. Two repetitions (chambers) were prepared for each time interval.

In each chamber 300 µL of cellular suspension ($5 \cdot 10^4$ cells/ml) were deposited in RPMI 1640 medium supplemented with 10 % FBS and 1 % P/S. The concentration of the culture was determined by an automated cell counting device (TC20™ Automated Cell Counter; BIO-RAD, ref.: 145-0101). The cells were incubated for 24 h (37 °C and 5 % CO_2) in order to guarantee their adherence to the surface of the chamber slide.

Next, the supernatant was removed and the chambers were washed with RPMI medium (free of serum) in order to eliminate any non-adhered cells. Into each of the chambers 300 µL of a

[4] The peritoneal cavity is the potential space between the membrane (peritoneum) that surrounds the abdominal wall and the membrane that surrounds the internal organs.

suspension of stationary phase promastigotes ($5 \cdot 10^5$ promastigotes/mL) were added in supplemented RPMI medium. The concentration of promastigotes was determined with a haemocytometer. The proportion of macrophages and promastigotes (1:10) and their respective concentrations were determined following Bilbao 2014. This ratio corresponds to the initial concentration of macrophages and the concentration of promastigotes inoculated. As macrophages were incubated for 24 h without inoculum and considering that they are a dividing population, it is not the ratio at the time of inoculation.

The chamber slides were incubated at 35 ^{o}C without CO_2 for 4 h in order to facilitate the infective conditions of the parasite. After this period, they were incubated at 37 ^{o}C and 5 % CO_2. After 24 h, the supernatant was removed and the chambers were washed twice with serum free RPMI medium in order to eliminate remaining promastigotes. After this, 300 µL of supplemented RPMI medium were added and the chamber slides were again incubated at 37 ^{o}C and 5 % CO_2.

Once the desired time interval had elapsed, the supernatant was removed and the chamber slides were left to dry at room temperature. In the cases of time intervals inferior to 24 h, two washes with RPMI medium were done. The slides were fixed with methanol and stained with 10 % Giemsa colouring during 10 minutes. The set up of the preparations was carried out with the set up medium DPX.

Finally, the number of amastigotes, healthy and infected macrophages were counted from a total of 300 cells in an optic microscope. The percentage of infection, the average number of amastigotes per infected macrophage and the parasitary index (IP) were calculated. The latter is calculated as the product of the other two parameters.

3. Description of the model

The standard description for IBMs is the ODD protocol as described in in Grimm et al. (2006). ODD stands for *Overview*, *Design concepts* and *Details* and it is described in detail in Annex A. The following is a briefer description of the most relevant aspects of the model.

3.1. General description

The aim of the model is to study and describe the *in vitro* infection of macrophages of the RAW 264.7 cell line by *L. infantum*. The dynamics of the community are not preconceived within the model but emerge from the interactions between individuals. This concept, emergence, has been discussed previously in Subsection 1.4.

The model considers two types of individuals or agents: macrophages and parasites. Macrophages can either be infected or healthy (i.e., not infected) and parasites can either be extracellular, adhered to the macrophages' external membrane or intracellular. Free parasites have the ability to move, adhere to macrophages and infect them. Parasites that have successfully infected a macrophage can reproduce within them. The individuals are characterised by several properties (Table 1), which are determined by a series of parameters (Table 2) that can be modified by the modeller. Many of the parameters that describe the behaviour of the individuals are difficult to determine experimentally, nor have they been found in bibliography. Instead, they have been adjusted by comparing simulation results to experimental results. This is known as parametrisation and will be described in detail in Section 4.

Table 1. Variables and parameters that characterise the properties of individuals.

	Name	Units	Description
Macrophage	diamac	µm	Diameter of macrophage
	edatmac	h	Age of macrophage
	imac		Identification index of macrophage
	estatmac		State of macrophage (0: healthy, 1: adhered or 2: infected)
	tinfect	h	Time the macrophage has been infected
	repromac		Indicates if the macrophage is able to reproduce (1: yes, 0: no)
	rlogisind	h^{-1}	R parameter of the logistic growth function of macrophages
Leishmania	diam	µm	Diameter of amastigotes
	edatam	h	Age of amastigotes
	estatam		State of amastigotes (0: free, 1: adhered or 2: intracellular)
	inmac		Index of the macrophage to which adhered and intracellular amastigotes are linked.

Table 2. Parameters that govern the behaviour of the individuals.

	Parameters	Units	Description	Source
Macrophages	nmacini		Initial number of macrophages	Experimental method
	diamacmax	µm	Maximum diameter of macrophages	Parametrisation
	rlogis	h^{-1}	Parameter R of the logistic function describing the growth of macrophages	Parametrisation
	tdmac	h	Duplication time of macrophages	(Assanga 2013)
	direpromac	µm	Minimum diameter for reproduction of macrophages	Parametrisation
	repart		Proportion of parasites divided between macrophages during duplication	Parametrisation
	maxinfect		Maximum number of parasites per infected macrophage	Parametrisation
Leishmania	namini		Initial number of parasites	Experimental method
	diammax	µm	Maximum diameter of parasites	(Gállego 2001)
	tmortam	h	Time a parasite is able to spend outside, without infecting a macrophage	Parametrisation
	distam	µm	Distance of movement of parasites	Parametrisation
	probenganxa		Probability of adhesion of a parasite to a macrophage	Parametrisation
	probinfecta		Probability of an adhered parasite of infecting a macrophage	Parametrisation
	tdam	h	Duplication time of parasites	Parametrisation

L. infantum is considered in promastigote or amastigote form depending on the stage of infection. The physiological differences are only taken into account in certain parameters that govern their behaviour (time of death, distance of movement), but the model does not consider different rules for the two forms. When inoculated, all parasites are in promastigote form. Once infection occurs, those parasites that have successfully infected macrophage transform to the amastigote form. The washing of the culture (see Section 3) is modelled by removing all free parasites (mostly promastigotes) at 24 h. The experimental removal is not complete; therefore, it is modelled in a way that most promastigotes are removed, but not with 100 % efficiency.

The space modelled is equivalent to the optic field of a microscope and consists of a grid of 400x400 spatial cells. The model runs for 72 time steps and each time step is equivalent to an hour. This time limit was chosen because, after 72 h, those cultures without medium renewal had a very poor final state. Those with the necessary medium renewal at 72 h showed a proliferation of cells that made the quantification of infection difficult and impractical.

There are many variables that are not and cannot be accounted for. The inclusion of stochasticity at every level is important for obtaining the appropriate emergent behaviours (Ferrer et al. 2010). This means that all processes are related to a certain probability of occurring and that all parameters are subjected to Gaussian noise with a standard deviation of 0.25.

3.2.Sequence of actions

In IBMs, actions, each governed by a rule, are executed in sequential order. The flow chart in Figure 9 and Figure 9 shows the sequence of actions that will presently be described. The chart is especially useful to visualise at what points the state of each individual is evaluated and how this state determines the course of actions it must follow. The code of the model can be found in Annex B.

The model, Figure 9, mirrors the experimental method, which means that first healthy macrophages are created according to their properties and distributed at random and without superposition in the modelled space. The inoculation of the promastigote culture, as described in Section 3, is considered as an input in the model. It involves the creation of the parasites separately from the macrophages, distributing them randomly in the available space.

Macrophages, as shown in Figure 9A, increment their age at each time step, which are equivalent to one hour. Given they have enough space, macrophages can grow following a logistic function depending on their diameter and the available surrounding space. Once they have reached sufficient diameter, they can duplicate by bipartition.

Parasites, as can be seen in Figure 9B, also increment their age with each time step. When they change state between extracellular, adhered and intracellular states, as described below, their age is set to zero and the count begins anew. The parameter and sequence of behaviour rules differ between the three states, so this is first evaluated to decide along which path each individual must proceed.

Extracellular and adhered parasites (those that have been unsuccessful in infecting a macrophage) have a certain probability of death related to their maximum age. Adhered parasites have double the maximum age with respect to free ones. At time step 24, all extracellular parasites die, without considering their age, because of the washing of the culture.

Extracellular parasites can also move a certain distance in a random direction. The distance of movement is reduced to half once 24 hours have passed. This is because all promastigotes have been removed and only free amastigotes remain. This difference in distance of movement is a way of taking into account the physiological differences between promastigotes and amastigotes.

Free parasites, both promastigotes and amastigotes, have a certain probability of adhering to macrophages they encounter and, once adhered, a certain probability of infecting them. Once the parasite is intracellular, it can reproduce, given it has attained a certain age, until the maximum number of parasites per macrophage is reached. Once this occurs, the macrophage burst, returning the parasites to the extracellular state.

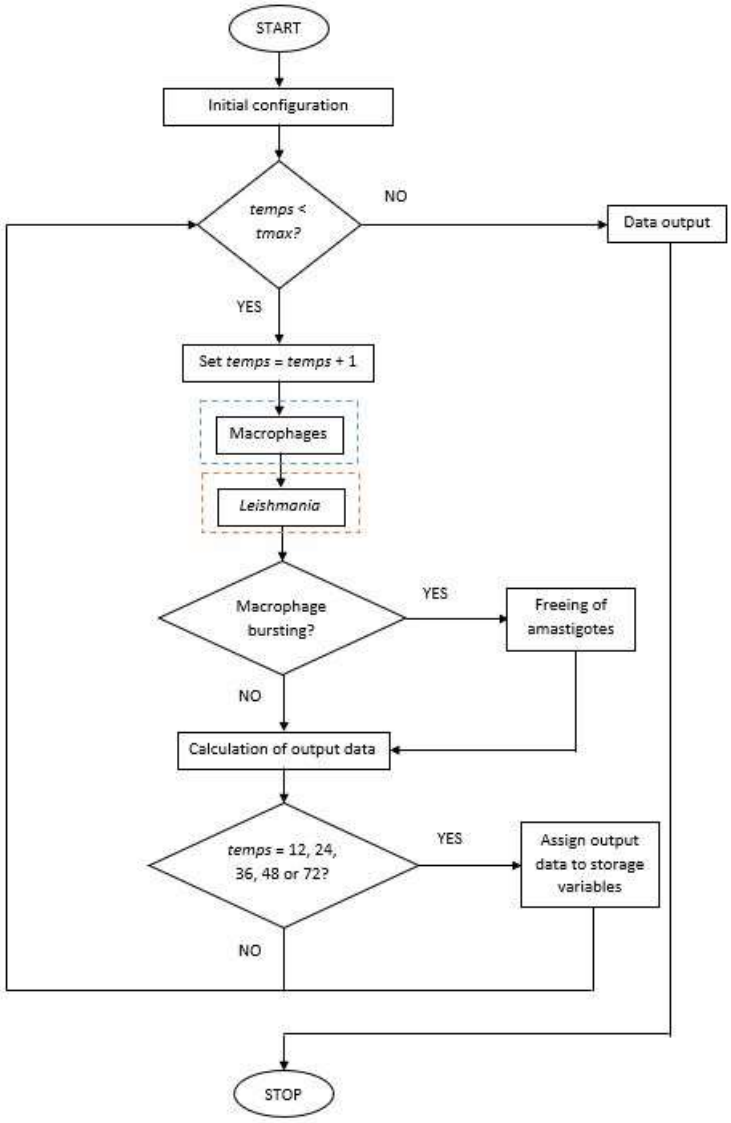


Figure 8. Flowchart of the macrophage - parasite model implemented in NetLogo.

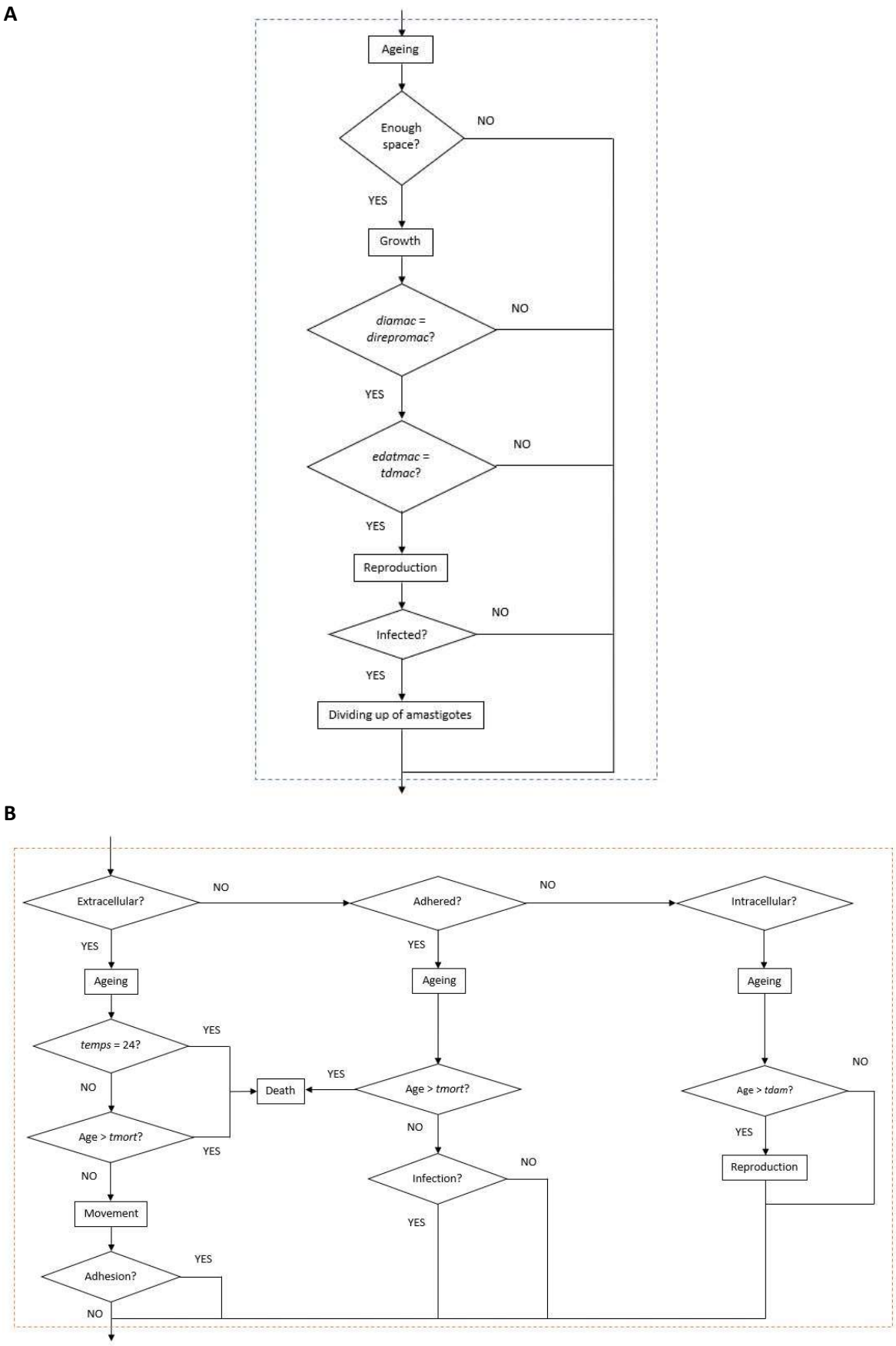


Figure 9. (A) Flowchart of the sequence of actions of macrophages and (B) flowchart for the sequence of actions of Leishmania.

3.3. Model implementation and simulation scheduling

The model is implemented in the NetLogo platform, the interface of the simulator is shown in Figure 10. The values of the parameters in Table 2 are stored in a commas-separated values (CSV) file. Each line of the file contains a parameter combination and has as many lines as parameter combinations wish to be tested.

The simulator reads the file one line at a time, assigning each value to the corresponding variable within the model. Because of the stochasticity at individual level, each simulation outcome is different, even for the same parameter combination. In order to obtain mean values for the output variables, the same combination is repeatedly simulated, as many times as the modeller deems appropriate. In this case, 30 repetitions were carried out.

The output variables are: percentage of infection, percentage of macrophages containing one or two amastigotes, percentage of macrophages containing three or four amastigotes and percentage of macrophages containing five or more amastigotes. The mean values for all the repetitions for these variables at time steps 12, 24, 36, 48 and 72 h are exported in a CSV file. Each line is the set of results corresponding to a parameter combination. This data is then processed with a different platform, in this case, MATLAB. The processing of data will be further described in Section 0.

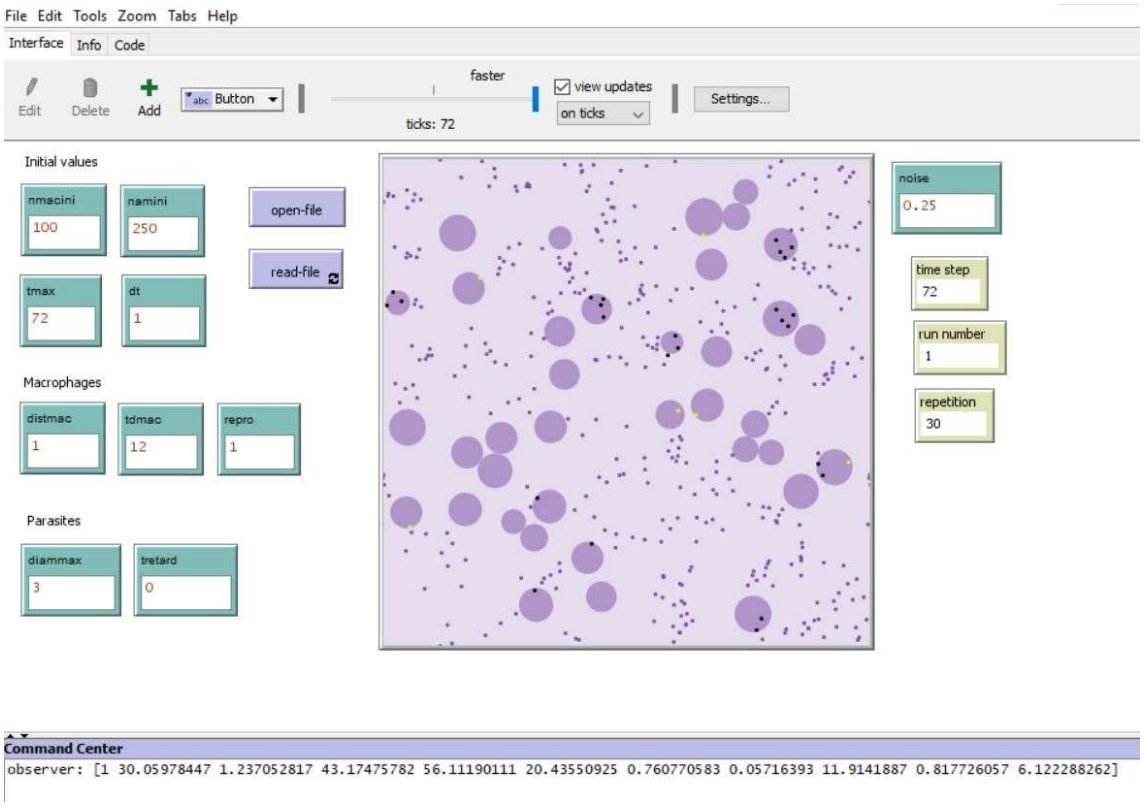


Figure 10. Interface of the NetLogo simulator (see text for details).

4. Analysis and parametrisation

Most of the parameters governing the behaviour rules of the individuals have not been determined experimentally. Therefore, they must be adjusted in a way that the simulated results resemble the experimental findings. This is not only a way to obtain a working model, it also contributes to shedding light on previously unknown values that are difficult to measure experimentally. The overall aim is to find the combination of parameter values that most resemble experimental findings and that have values representative of the proposed behaviour mechanisms.

The course of action to be followed, as will be described shortly, is shown in Figure 11. It starts with values adjusted in Raventós (2017). The parameters are explored one by one and the discrepancy between experimental and simulated results is assessed. An interval where the discrepancy is minimum is defined for further exploration. These intervals make up the parameter space, the set of all the possible combinations of parameter values contained in the model. The parameter space is explored and the result is a distance between experimental measurements and simulation outcomes for each parameter combination. One the one hand, the combination with the minimum distance value is found (initial optimum values). On the other hand, the results of the exploration permit the fitting of an n-dimensional curve to the parameter space. In this case, 10 parameters are considered, so the curve would have 10 dimensions. By studying this curve, an optimum parameter combination can be found.

In this bachelor thesis no curve has yet been adjusted. The parametrisation process has only reached the stage of obtaining the results of the parameter space exploration and finding the initial optimum value.

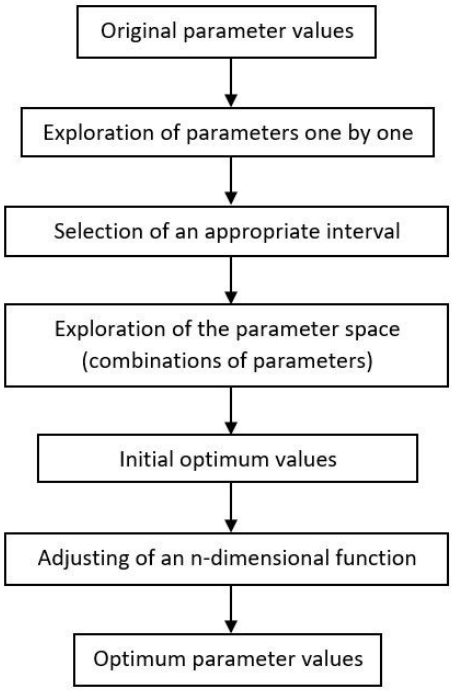


Figure 11. Diagram of the steps of the parametrisation process.

4.1. Definition of mathematical distance

A mathematical distance is defined in Equation (1) as a measure of the discrepancy between certain experimental results and the corresponding simulation outcomes.

$$d = \sqrt{\sum_{i=1}^{n} \left(x_{exp,i} - x_{sim,i}\right)^2} \tag{1}$$

This distance is used for evaluating the discrepancy between the experimental and simulated systems. The variables used are:

- Infection percentage: the percentage of infected macrophages among the total number of macrophages.
- Percentage of macrophages with 1 or 2 parasites within them.
- Percentage of macrophages with 3 or 4 parasites within them.
- Percentage of macrophages with 5 or more parasites within them.

The distances are evaluated at the time-points when experimental measurements are available (12, 24, 36, 48 and 72 hours). The distance that assesses the difference of percentage of infection is known as the infection distance (d_{inf}). The distances that evaluate the discrepancy among the number of parasites per macrophage are denoted d_{1+2}, d_{3+4} and $d_{\geq 5}$, respectively. Together, these last three distances make up the distribution distance (d_{distr}), as seen in Equation (2). Finally, d_{inf} and d_{distr} are combined to form a global distance (d_{global}), as in Equation (3). The global distance is used to evaluate the discrepancy between experimental results and the simulation output.

$$d_{distr} = \sqrt{d_{1+2}^2 + d_{3+4}^2 + d_{\geq 5}^2} \tag{2}$$

$$d_{global} = \sqrt{d_{inf}^2 + d_{distr}^2} \tag{3}$$

Different combinations of parameters will be explored, as described in the following sections. For each of the considered parameter combinations, 30 simulations are run. As a result, 30 distances are calculated for each of the five intervals evaluated (12, 24, 36, 48 and 72 hours). In total, this accounts for 150 distances. First, the mean distance for the five time-points is calculated. The result is a mean infection distance and a mean distribution distance for each simulation. With these values, the global distance is calculated for each simulation. Finally, a mean for the 30 simulations of each parameter combination is calculated. The final result is a single global distance value for each parameter combination.

The simulations are programmed with NetLogo and the calculation of distances is carried out using MATLAB. The process is described in detail in the following sections.

4.2. Preliminary analysis

In Raventós (2017), the parameters where adjusted one by one, i.e., by testing a range of values for one parameter while keeping the rest constant. Parameters whose values could be found in literature or pertained to the experimental method, listed in Table 3, were not tested. The parameters that were adjusted are detailed in Table 4. For the preliminary analysis, the code used to calculate the mathematical distance from the simulation output and the experimental results can be found in Annex C.

Table 3. Name, description and values of parameters governing the behaviour of macrophages and parasites found in literature or determined by experimental method.

	Parameter	Description	Value	Units	Source
Macrophages	Nmacini	Initial macrophages	100	Individuals	Experimental method
	Tdmac	Macrophage duplication time	12	h	(Assanga 2013)
Leishmania	Namini	Initial parasites	250	Individuals	Experimental method
	Diammax	Maximum parasite diameter	3	µm	(Gállego 2001)

Table 4. Name, description and values of the parameters governing the behaviour of macrophages and parasites not found in bibliography and adjusted in Raventós (2017).

	Parameters	Description	Units	Original value
Macrophages	Diamacmax	Maximum diameter	µm	30
	Rlogis	R of logistic function	h^{-1}	1.5
	Direpromac	Minimum diameter for reproduction	µm	60
	Repart	Proportion of distribution during macrophage duplication		0.6
	Maxinfect	Maximum of parasites per infected macrophage	am/mac	6
Leishmania	Tmortam	Mortality	h	36
	Distam	Distance of movement	µm	15
	Probenganxa	Adhesion probability		1
	Probinfecta	Infection probability		0.04
	Tdam	Duplication time	h	6

The next step is to study the parameters in combination. In order to do this, we must define a parameter space. The aim of the preliminary analysis is to explore a series of parameter values one by one in order to find an interval where the mathematical distance between experimental results and simulation outcomes is minimum as shown in Figure 12. Each of the points is the mean value of the mathematical distance of thirty simulations, assessed as described above. This is necessary to ensure that an appropriate parameter space is chosen for subsequent study.

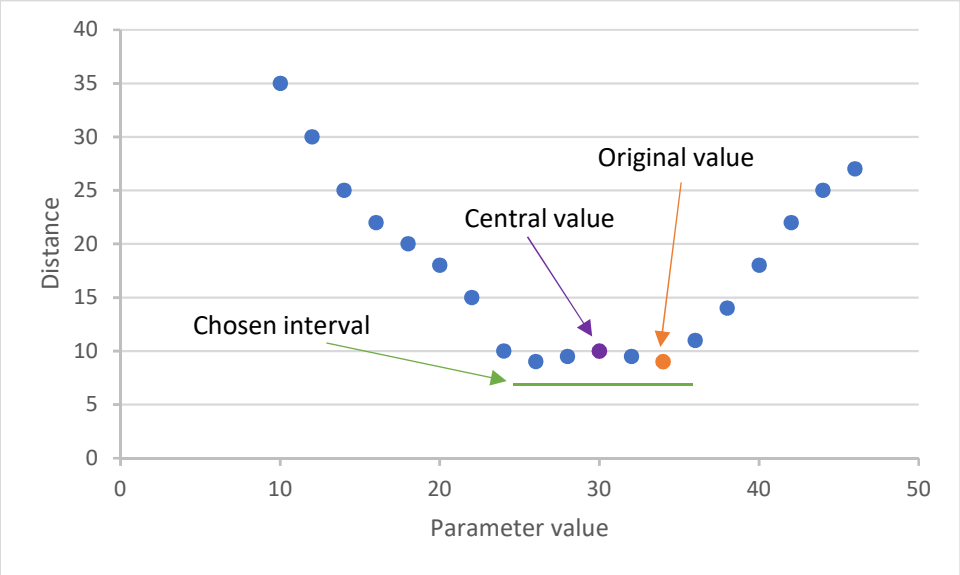


Figure 12. Hypothetical result of the preliminary analysis, where each of the point is the mean value of the global mathematical distance of thirty simulations.

The interval is defined by taking a value more or less in the centre of it (central value), as shown in Figure 12. Hypothetical result of the preliminary analysis, where each of the point is the mean value of the global mathematical distance of thirty simulations., and defining upper and lower bounds, as in Equations (4) and (5):

$$\text{Upper bound} = 1.2 \cdot \text{Central value} \tag{4}$$

$$\text{Lower bound} = 0.8 \cdot \text{Central value} \tag{5}$$

One must bear in mind that Figure 12 does not represent real data and that it is used merely for the purpose of transmitting information. This type of figure will be encountered further along in the thesis, when the results of this preliminary analysis are presented in Section 5.2. These figures are not used to extract precise results. Their purpose is to have an overview of the behaviour of each parameter. With this knowledge, a convenient region of parameter values can be chosen for further study.

4.3. Exploration of the parameter space

In this section, the parametrisation process developed will be described in detail, as shown in Figure 13. Parameter combinations from the parameter space defined above will be sampled and stored in a CSV file. A series of simulations will be performed with these combinations in the NetLogo platform. The results of these will again be stored in a CSV file and later analysed with MATLAB.

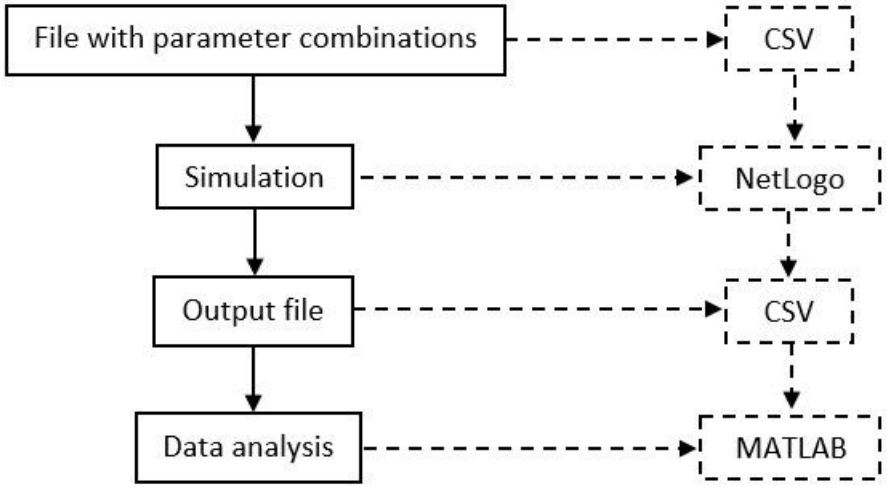


Figure 13. Flow chart of the steps that make up the parametrisation process, with the formats and platforms used for each step.

4.3.1. Latin hypercube sampling

Because it is not feasible to simulate all the possible combinations, an adequate number of samples have to be taken from the parameter space. These have to be taken with care in order to get a good representation of the sample space with a limited amount of simulations (Mckay et al. 2000).

The technique used is known as Latin Hypercube Sampling (LHS). It consists of partitioning the sample space and ensures that all portions are samples, i.e., that each input variable has all portions of its distribution represented by input values (Iman et al. 1981). The LHS was carried out with the MATLAB function *lhsdesign*, as can be seen in the code in Annex D. The intervals used are the ones resulting from the Preliminary Analysis and 1000 samples were taken. The resulting matrix is stored in CSV format with semicolon separators.

Each line of the file contained a parameter combination for the 10 parameters described above, each combination a sample of the parameter space. Each line contains 11 values. The first an identifier for the combination, and its value ranges from 1 to the number of samples taken from the parameter space, 1000 in this case. The other ten numbers are the corresponding parameter values.

4.3.2. Incorporation into the model

Figure 14 shows the flowchart of the code implemented in NetLogo (code in Annex A) in order to integrate the proposed parametrisation methodology and the parasite – macrophage IBM. Figure 14 shows the overall structure of the code and Figure 15 shows specifically the model. Note that there are some structural changes with respect to the flowchart presented in Figure 9. Figure 9B and 9C remain the same. The differences do not affect the working of the model, but the input and output of data.

The NetLogo CSV extension was used in order to be able to read the file where the parameters where stored and import them to the simulator. The input CSV file contains as many parameter combinations as samples taken from the parameter space. The first number of the line is the identifier, followed by the parameter values. At each run, a line is read and each value assigned to its corresponding variable within the model.

For each parameter combination 30 simulations were carried out, each with its own setup routine, in order to start each simulation afresh. This number of repetitions was chosen in order to be able to compute mean values of the output variables considering the stochasticity of the model.

For each combination of parameters, the simulator returns four output variables explained above at the five time-points mentioned (12, 24, 36, 48 and 72 h); a total of 20 values. These data are stored in another CSV file, where each line contains the simulation outcome of a parameter combination, the first number being again the identifier of the combination.

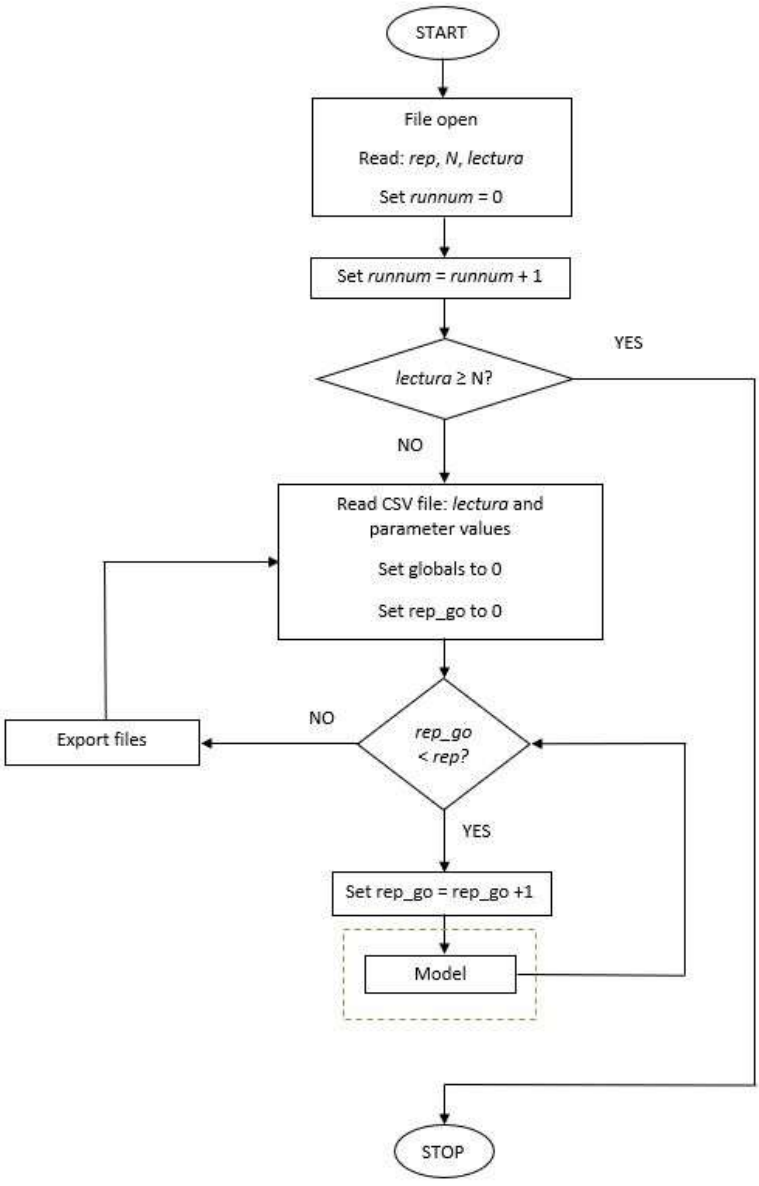


Figure 14. Flowchart of the parametrisation methodology implemented in NetLogo.

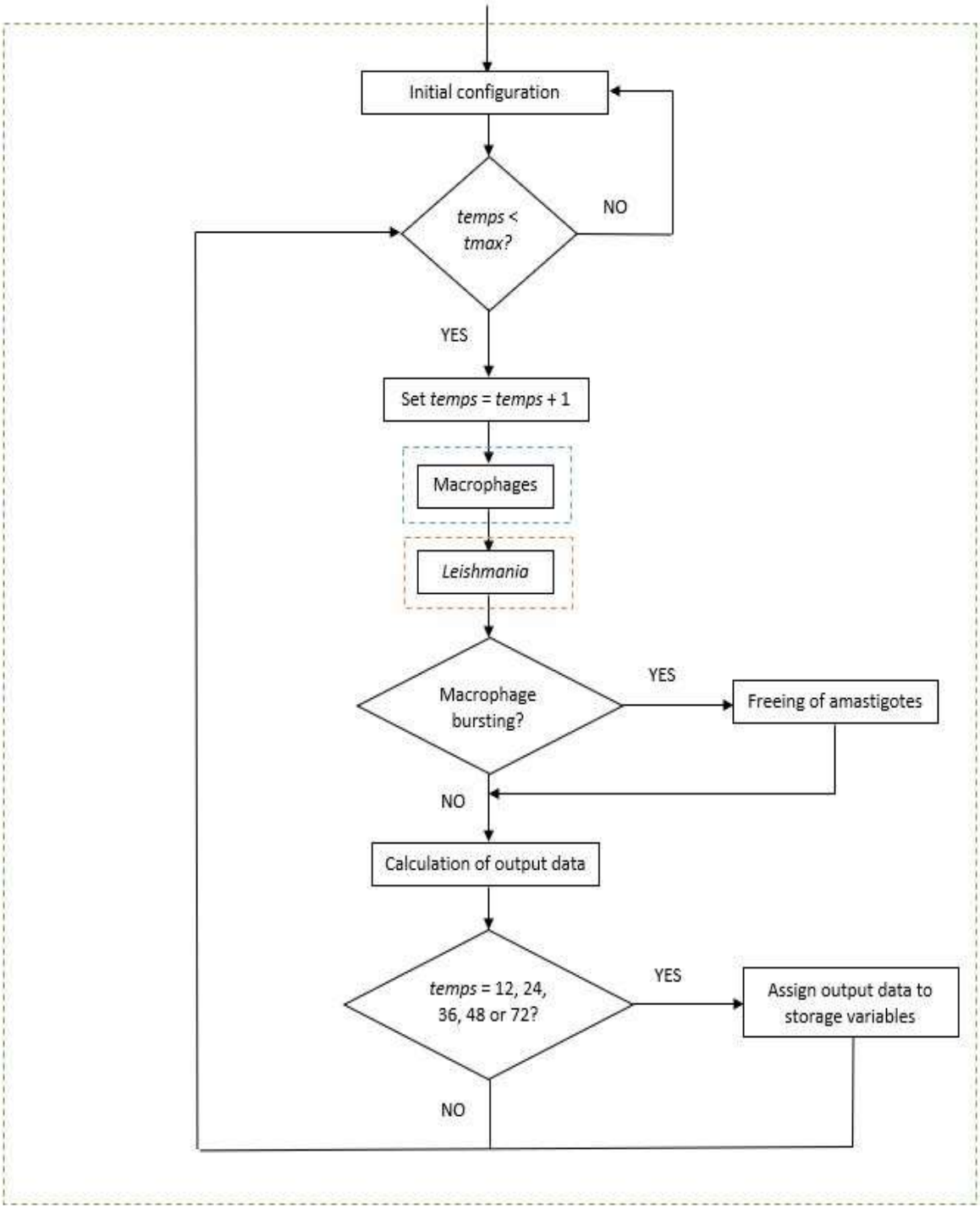


Figure 15. Flowchart of the parasite - macrophage model modified to fit the parametrisation methodology.

4.3.3. Analysis of the results

The CSV with the outcomes of the simulations are analysed with MATLAB. This analysis consists of the evaluation of the resulting distance for each of the individual runs, and then computing the mean distance value for each parameter combination. The process is detailed with the code and flowchart in Annex E.

The output is an Excel file that contains, in each line, the identifier of the parameter combination, the values of the parameters and the mathematical distance associated to that parameter combination. This distance is the result of comparing the experimental results to the simulation output of that particular parameter combination.

5. Results

5.1. Experimental results

Table 5, Figure 16 and Figure 17 show the experimental results obtained from the assays performed by Raventós (2017). The analysed variables were the percentage of infected macrophages (Figure 16) and the parasite load of these macrophages (Figure 17).

In Figure 16, the results show an initial increase in the percentage of infected macrophages, followed by a decrease because of their lysis. The peak is reached at 36 h. In Figure 17, the largest subpopulation is that of macrophages containing 1 or 2 amastigotes. The other two subpopulations, macrophages containing 3 or 4 amastigotes and macrophages containing more than 5 amastigotes, have similar sizes all along the evolution of the culture. All three subpopulations show peaks at 36 h.

Table 5. Experimental results for percentage of infection, macrophages with 1 or 2 amastigotes, macrophages with 3 or 4 amastigotes and macrophages with 5 or more amastigotes, at five time-points.

Time	% Infection	1 or 2 amastigotes	3 or 4 amastigotes	5 or more amastigotes
12	22.25	19.5	2.5	0.25
24	37.17	25	6.5	5.67
36	50	29	11	10
48	30.33	21.83	5.67	2.83
72	18.33	17.67	0.67	0

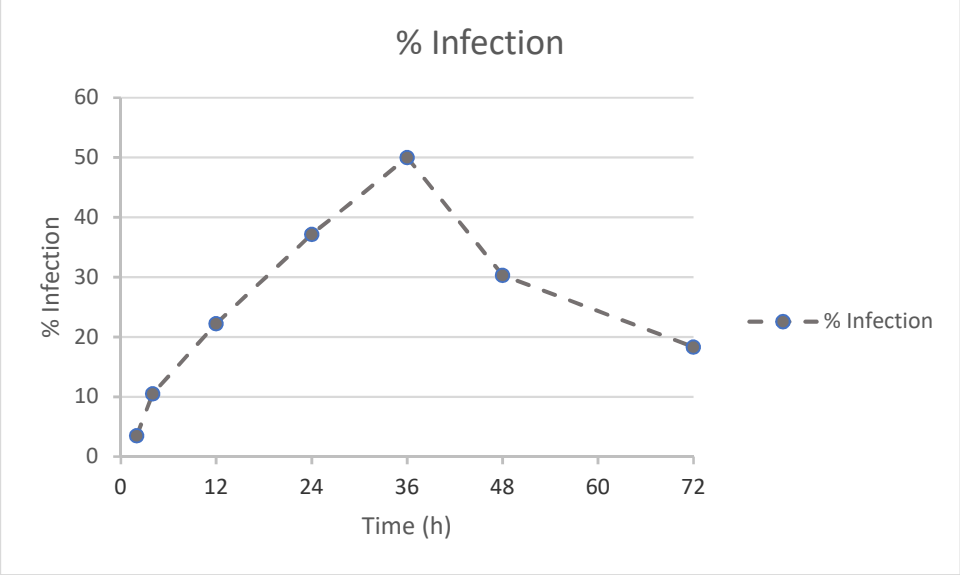


Figure 16. Experimental results of the percentage of infection.

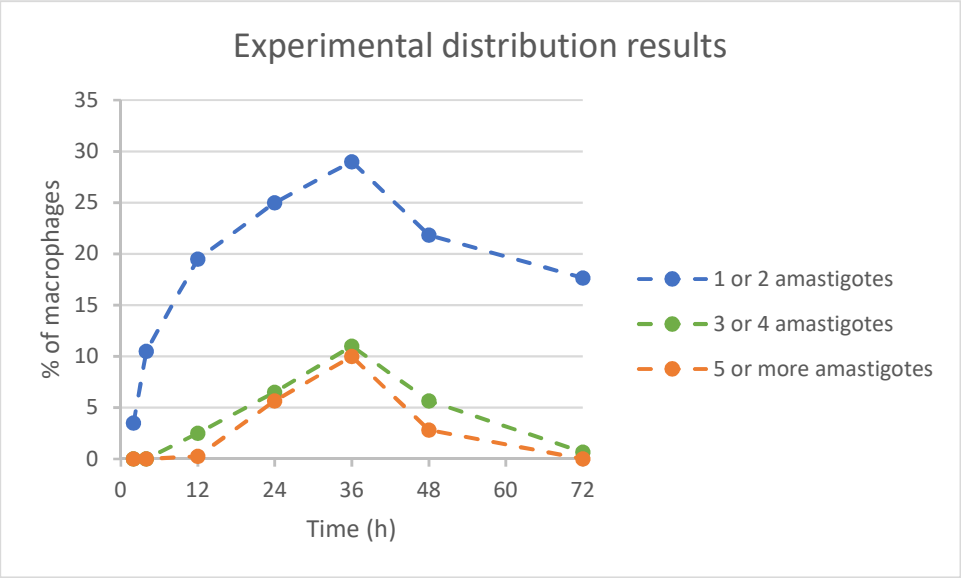


Figure 17. Experimental results of the distribution of amastigotes within macrophages.

To give an idea of the uncertainty of these values, the results for each of the two repetitions, as obtained in Raventós 2017, are shown in Figures Figure 18*Figure* 21. There was only one successful repetition for 72 h. The relative position between the points is an indicator of the uncertainty of the measurements. Most points show high similitude between repetitions, but the measurements at the time point where each curve reaches its peak shows higher discrepancy than the rest.

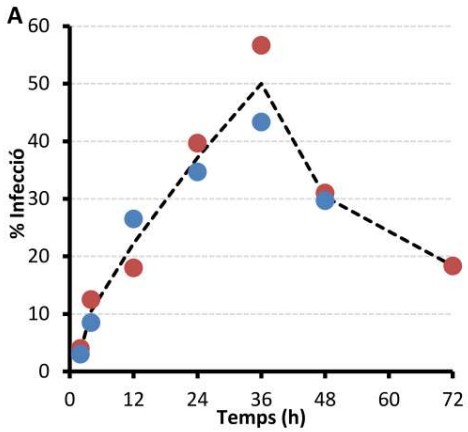


Figure 18. Experimental results of the evolution of percentage of infection for two repetitions. There was only one successful repetition for 72h. Source: Raventós 2017.

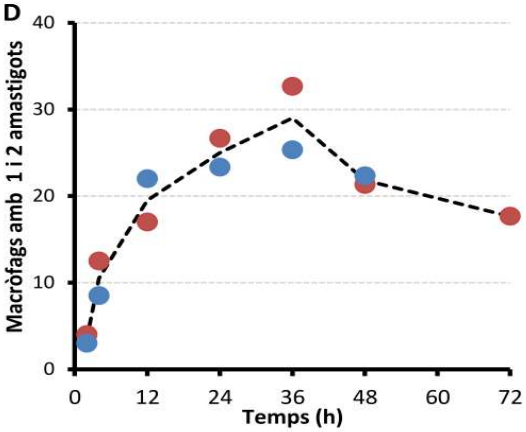


Figure 19. Experimental results of the evolution of the percentage of macrophages with 1 or 2 amastigotes for two repetitions (except 72 h). Source: Raventós 2017.

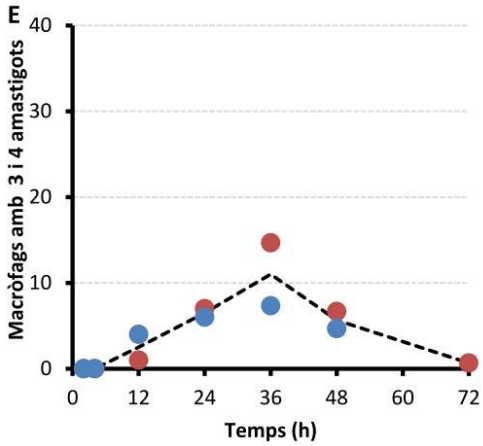


Figure 20. Experimental results for the evolution of the percentage of macrophages with 3 or 4 amastigotes for two repetitions (except 72 h). Source: Raventós 2017.

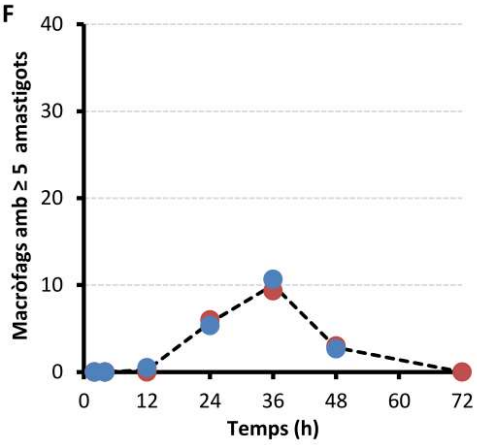


Figure 21. Experimental results for the evolution of the percentage of macrophages with 5 or more amastigotes for two repetitions (except 72 h). Source: Raventós 2017.

5.2. Results of preliminary analysis

In the following, Figure 22 to Figure 31 collect the results of the preliminary analysis of parameters. The horizontal axis represents the range of parameter values and the vertical axis, the global mathematical distance between the experimental measurements and the corresponding simulation outcomes, as calculated as explained in Section 4.1. The plots help visualise which values show a closer correspondence between experimental results and simulation results.

All parameters, except *rlogis*, present a region in their distribution where the mathematical distance between experimental and simulated values is minimal. This indicates that there is a region of parameter values that is more suitable for the correct functioning of the model. It also indicates that the value of these parameters has an impact on the result of the simulation, i.e., that the model is sensitive to them. This is not the case for *rlogis*, which indicates that its value has a negligible impact on the simulation outcome. Note also, that beyond a certain point, the value of parasite mortality (*tmortam*) is no longer transcendent to the outcome of the simulation.

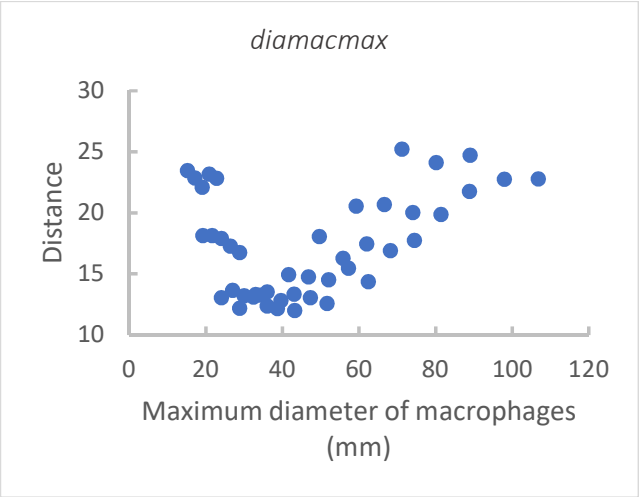


Figure 22. Result of the preliminary analysis of the parameter diamacmax*, the maximum diameter of macrophages (µm).*

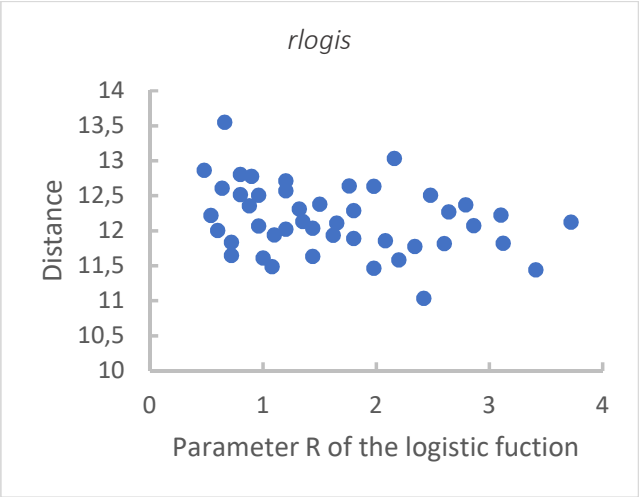


Figure 23. Results of the preliminary analysis of the parameter rlogis*, the r parameter of the logistic function (h^{-1}).*

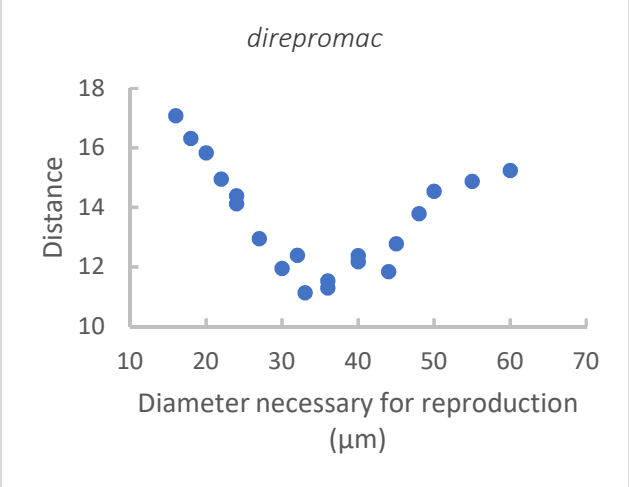


Figure 24. Results of the preliminary analysis of the parameter direpromac*, the minimum diameter necessary for the duplication of macrophages (µm).*

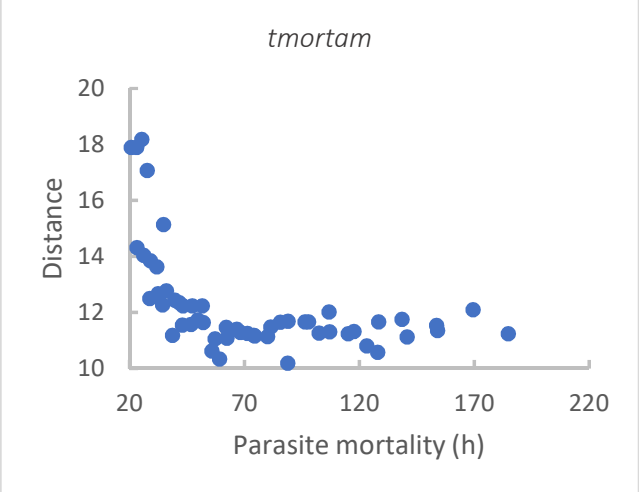


Figure 25. Results of the preliminary analysis of the parameter tmortam*, the mortality time for parasites (h).*

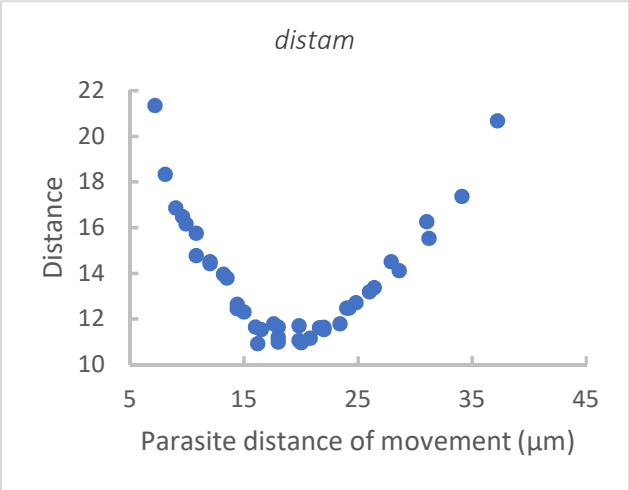


Figure 26. Results of the preliminary analysis of the parameter distam, the distance of movement of free parasites (µm).

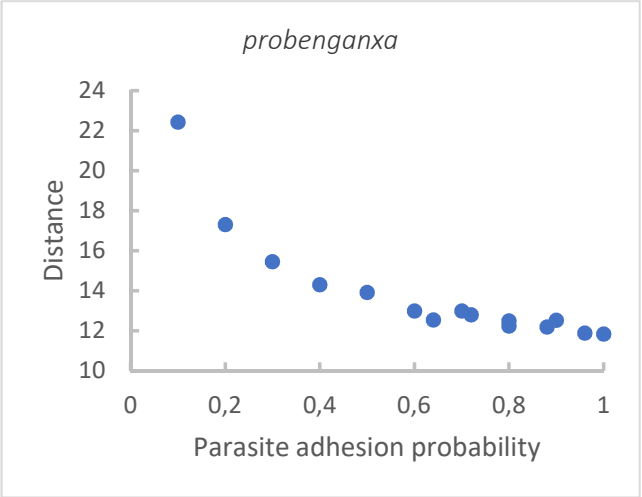


Figure 27. Results of the preliminary analysis of the parameter probenganxa*, the probability a free parasite of adhering to a macrophage.*

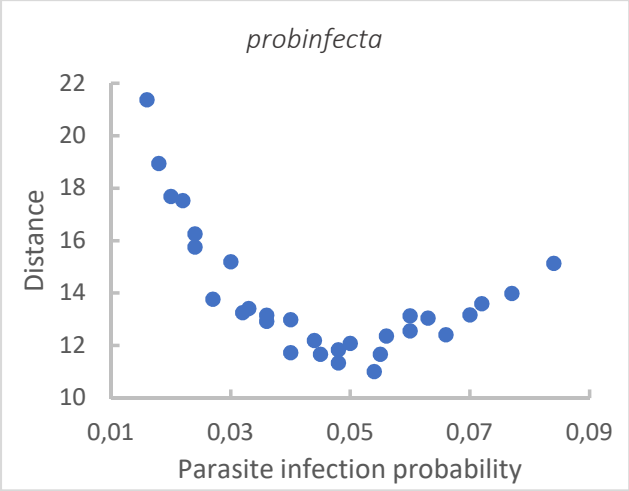


Figure 28. Results of the preliminary analysis of the parameter probinfecta*, the probability of an adhered parasite of infecting a macrophage*

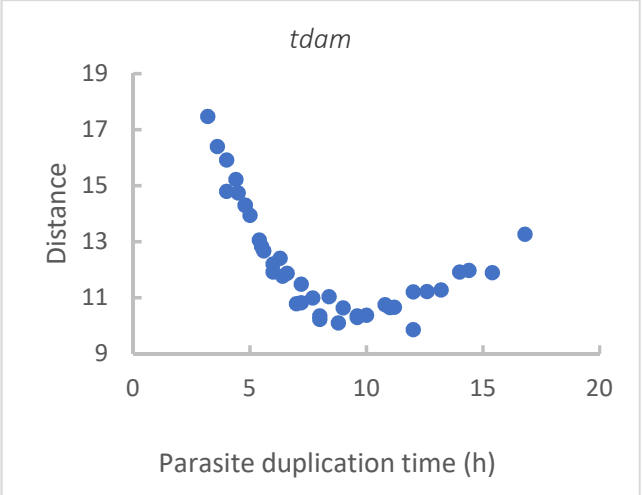


Figure 29. Results of the preliminary analysis of the parameter tdam*, the duplication time of amastigotes (h).*

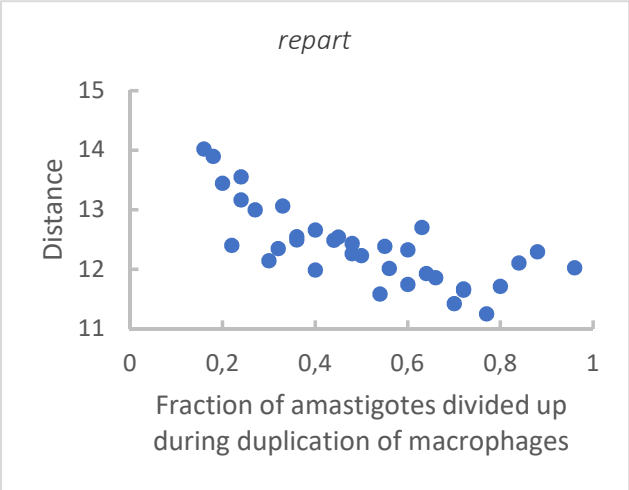


Figure 30. Results of the preliminary analysis of the parameter repart*, the fraction of amastigotes divided up during the duplication of macrophages.*

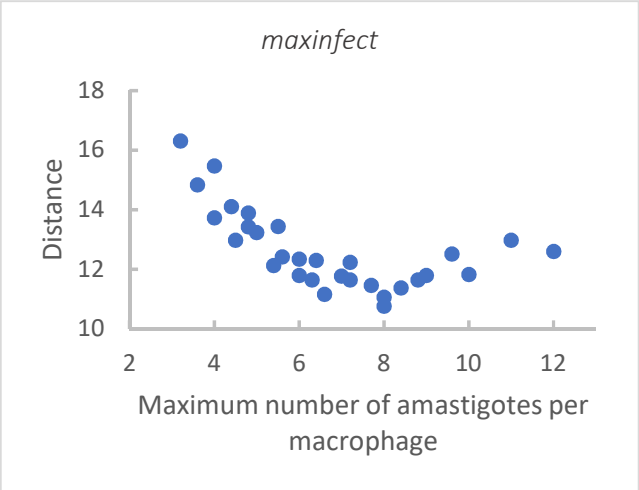


Figure 31. Results of the preliminary analysis of the parameter maxinfect*, the maximum number of amastigotes per macrophage.*

Table 6. Results of the preliminary analysis of parameters: central value, upper and lower bounds of the interval of parameter values where the distance between experimental and simulation results is minimum.

	Parameters	Description	Units	Lower bound	Central value	Upper bound
Macrophages	Diamacmax	Maximum diameter	µm	28.8	36	43.2
	Rlogis	R of logistic function	h^{-1}	1.2	1.5	1.8
	Direpromac	Minimum diameter for reproduction	µm	28.8	36	43.2
Leishmania	Tmortam	Mortality	h	41.6	52	62.4
	Distam	Distance of movement	µm	16	20	24
	Probenganxa	Adhesion probability		0.64	0.8	0.96
	Probinfecta	Infection probability		0.04	0.05	0.06
	Tdam	Duplication time	h	8	10	12
	Repart	Proportion of distribution during macrophage duplication		0.56	0.7	0.84
	Maxinfect	Maximum of parasites per infected macrophage	am/mac	5.6	7	8.4

5.3. Results of exploring the parameter space

The parameter combination that best adjusts to the experimental results, after the LHS exploration of the parameter space, can be found in Table 7. Figure 32 and Figure 33 show comparisons of the experimental results and the simulation outcome for this parameter combination. The mathematical distance value for this parameter combination is lower than that of the initial parametrisation carried out in Raventós (2017); 8.38 and 13.88, respectively . This means that, overall, the model fits the experimental system better with this parameter combination.

The adjusted values offer a biological interpretation that gives information on the behaviour of the individuals. The value of the parameter distam (17.4 µm), the distance the parasite moves in the extracellular medium, is relatively small when compared to the diameter of the macrophage (30.6 µm), especially considering that the amastigote only moves a distance half this value (8.7 µm). The values of probenganxa (0.7) and probinfecta (0.05) show that when a parasite comes in contact with a macrophage it almost always adheres successfully (approximately 70 % of the time). However, this does not mean that the parasite will be successful in infecting the macrophage, which happens around 5 % of the time.

The most interesting parameter is repart. Its value, 0.64, indicates that when an infected macrophage divides by bipartition, 64 % of the parasites within this macrophage go to one of the daughter cells and the other 36 %, to the other. On the one hand, questions about the structure and working of the parasite vacuole arise: does it split? On the other hand, it could be an indication that macrophages are often infected by more than one parasite and, therefore, have more than one parasite vacuole, which are distributed into different daughter cells at division.

Table 7. Parameter combination with least discrepancy between experimental results and simulation outcome.

	Parameter	Units	Description	Adjusted value
Macrophages	diamacmax	µm	Maximum diameter of macrophages	30.63
	rlogis	h^{-1}	Parameter R of the logistic function describing the growth of macrophages	1.42
	direpromac		Minimum diameter for reproduction of macrophages	37.36
	repart		Proportion of parasites divided between macrophages during duplication	0.64
	maxinfect		Maximum number of parasites per infected macrophage	6.33
Leishmania	tmortam	h	Time a parasite is able to spend outside, without infecting a macrophage	44.34
	distam	µm	Distance of movement of parasites	17.39
	probenganxa		Probability of adhesion of a parasite to a macrophage	0.699
	probinfecta		Probability of an adhered parasite of infecting a macrophage	0.052
	tdam	h	Duplication time of parasites	8.43

In Figure 32, the experimental results and simulation outcome of the percentage of infection are compared. The simulation outcome shows an appropriate behaviour, approximately following the shape of the experimental curve. Both curves show a peak at the same time-point, but the individual points show some discrepancy.

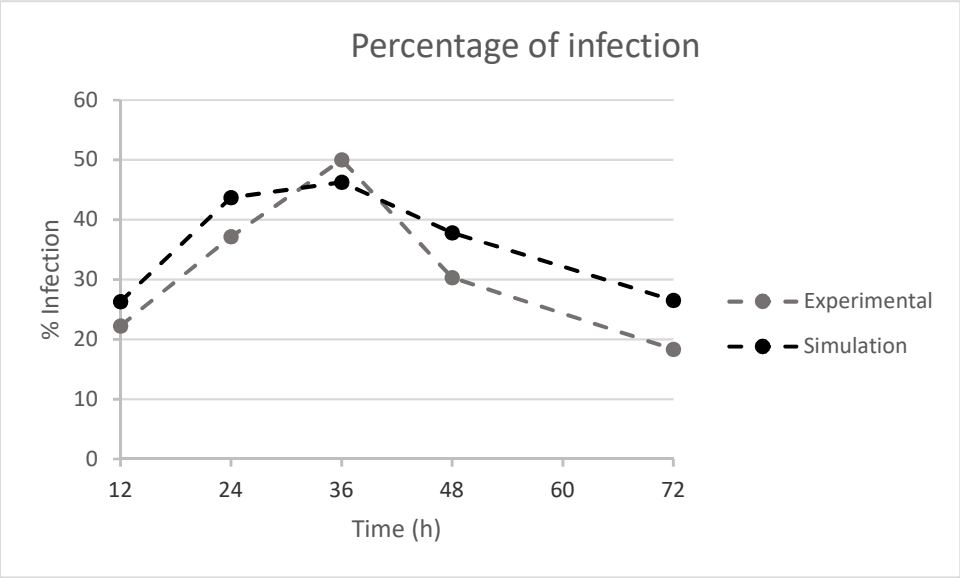
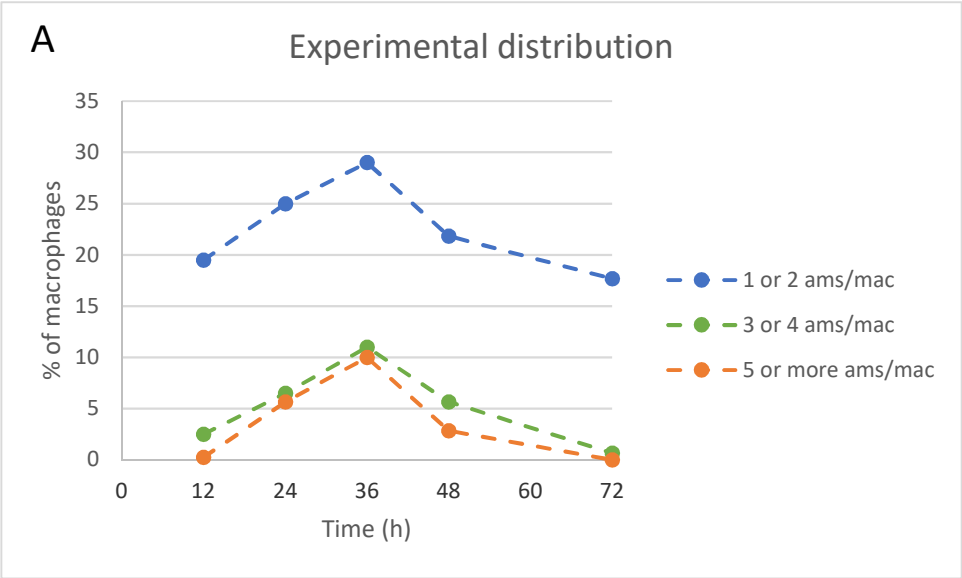


Figure 32. Comparison on the percentage of infection between the experimental results and the simulation outcome.

Figure 33A shows the distribution of parasites within macrophages for the experimental results and Figure 33B, the simulation outcome. The shapes of the curves are similar and so is their hierarchy. On the one hand, the simulation curves for percentage of macrophages with 3 or 4 amastigotes and with 5 or more amastigotes (green and orange) present peaks at the same time-points as the experimental curves. However, their end value differs noticeably. On the other hand, the experimental and simulation curves for the percentage of macrophages with 1 or 2 amastigotes (blue) present peaks at different time-points. However, their final maximum values are similar, as are their final values.



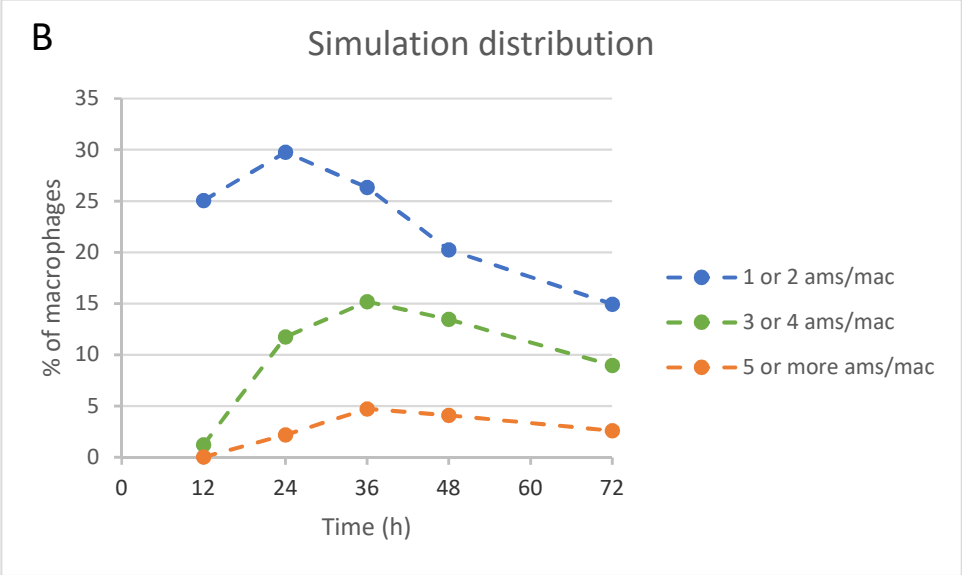


Figure 33. (A) Experimental results of the distribution of parasites within macrophages and (B) simulation outcome of the distribution of parasites within macrophages.

6. Discussion and conclusions

When studying the behaviour of complex systems such as a parasite – host cell culture is difficult to determine how the elements and their interactions contribute to the overall behaviour of the system. IBMs are a useful tool for studying complex systems because they focus on individual behaviour and interactions. The rules governing the model are those of individual agents, from which the macroscopic behaviour of the population then emerges.

One of the most difficult steps when building IBMs is their parameterization. On the one hand, these models usually involve a large number of parameters that must be determined. On the other hand, the kind of data needed for univocally determining their values is not always easy to be determined experimentally. Therefore, a parameterization process to obtain the best fit is always required, and this process must be optimised in order to explore possible combinations and select the best solution.

In this bachelor thesis, a mathematical methodology for the process of parametrisation is presented. It optimises the process of parametrisation in terms of results, speed and practicality for the user. The criterion used, mathematical distance, can be calculated from any experimental results and simulation outcome, independently from the variable measured. This means that the methodology is extensible other types of parasite cultures: different strains, species and mediums. For example, to assess the effects of drugs by studying the changes in parameter values. A qualitative and quantitative interpretation of the effects of potential drugs could give information on the mechanism of action of these compounds. It could even be extended to other parasite models such as malaria and Chagas' disease.

The development of an IBM that simulates *L. infantum* infection of RAW 264.7 cells sheds light on the behaviour of both types of individuals, including the parasites mechanism of action, increasing the understanding of this particular culture. It also allows the estimation of parameters that are difficult to determine experimentally. For example, the time a parasite can survive in the extracellular state is approximately 44 h and the maximum number of parasites a macrophage is capable of holding is around 6.

The model has been developed using the NetLogo simulation platform and its outcome has been analysed by means of MATLAB. The two functions, simulation and analysis, could be integrated and simulation speed accelerated by using a different programming language, for example C. This integration would make the simulations much faster and, especially, would allow the storage of much more detailed data in data arrays, which is very complicated in NetLogo. The complexity of both the model and the parametrisation process demand a more powerful platform in order to be able to increase the accuracy and quality of the results.

It is important to stress that work such as this is essentially interdisciplinary. Some aspects require expertise in parasitology, others, in modelling. The most important aspect, though, is the dialogue between the parasitologists and modellers. On the one hand, parasitologists must optimise the experiments and measurements in order to facilitate the comparison between the model and the experimental system. On the other hand, modellers must integrate the experimental results into the model, both in the behaviour rules and in the parameter values. Thus, this dialogue allows the further development and optimisation of the experimental design and the IBM.

Returning to Figure 7, which shows the stages in the development of a model and its corresponding experimental design, the steps can be retraced. In Raventós (2017), the

experimental method was developed and an initial parametrisation of the model was done. In this bachelor thesis, the initial evaluation, the determination of the parameter space and the exploration of this parameter space were carried out. The next steps include fitting an n-dimensional function to the parameter space and a global analysis of its outcome. The result of this will allow the further development and optimisation of the experimental design and the IBM. What is clear is that is first parametrisation is not nearly enough and that the cycle in Figure 7 needs to be repeated many times in order to reach a satisfactory result.

7. References

Alvar, J., Yactayo, S. & Bern, C., 2006. Leishmaniasis and poverty. *Trends in Parasitology*, 22(12), pp.552–557.

Assanga, I., 2013. Cell growth curves for different cell lines and their relationship with biological activities. *International Journal of Biotechnology and Molecular Biology Research*, 4(4), pp.60–70. Available at: http://academicjournals.org/journal/IJBMBR/article-abstract/6979C3B2546.

Aulner, N. et al., 2013. High Content Analysis of Primary Macrophages Hosting Proliferating Leishmania Amastigotes: Application to Anti-leishmanial Drug Discovery A. R. Satoskar, ed. *PLoS Neglected Tropical Diseases*, 7(4), p.e2154. Available at: http://dx.plos.org/10.1371/journal.pntd.0002154 [Accessed May 27, 2018].

Avila, J.L. & Casanova, M.A., 1982. Comparative Effects of 4-Aminopyrazolopyrimidine, Its 2'-Deoxyriboside Derivative, and Allopurinol on In Vitro Growth of American Leishmania Species. , 22(3), pp.380–385. Available at: http://aac.asm.org/content/22/3/380.full.pdf [Accessed March 28, 2018].

Beach, D.H., Goad, L.J. & Holz, G.G., 1988. Effects of antimycotic azoles on growth and sterol biosynthesis of Leishmania promastigotes. *Molecular and biochemical parasitology*, 31(2), pp.149–62. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2847043 [Accessed March 28, 2018].

Beach, D.H., Holz, G.G. & Anekwe, G.E., 1979. Lipids of Leishmania promastigotes. *The Journal of parasitology*, 65(2), pp.201–16. Available at: http://www.ncbi.nlm.nih.gov/pubmed/448607 [Accessed March 28, 2018].

Berman, J.D. et al., 1989. Biochemistry of Pentostam resistant Leishmania. *The American journal of tropical medicine and hygiene*, 40(2), pp.159–64. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2537584 [Accessed March 28, 2018].

Berman, J.D., Gallalee, J. V & Hansen, B.D., 1987. Leishmania mexicana: uptake of sodium stibogluconate (Pentostam) and pentamidine by parasite and macrophages. *Experimental parasitology*, 64(1), pp.127–31. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3038590 [Accessed March 28, 2018].

Bhuniya, D. et al., 2015. Aminothiazoles: Hit to lead development to identify antileishmanial agents. *European Journal of Medicinal Chemistry*, 102, pp.582–593. Available at: https://ac-els-cdn-com.recursos.biblioteca.upc.edu/S0223523415301975/1-s2.0-S0223523415301975-main.pdf?_tid=263fdf92-e135-4b09-963b-a0bd80809ddb&acdnat=1520634099_68f98d3cd46bd2580f736d30f89c5da4 [Accessed March 9, 2018].

Bilbao Ramos, P.E., 2014. *Caracterización molecular y evaluación de la susceptibilidad de especies de Leishmania, circulantes en Bolivia, a nuevos fármacos y formulaciones*. Universidad Complutense de Madrid. Available at: http://eprints.ucm.es/28985/ [Accessed April 5, 2018].

Bouazizi-Ben Messaoud, H. et al., 2017. Changes in Lipid and Fatty Acid Composition During Intramacrophagic Transformation of Leishmania donovani Complex Promastigotes into Amastigotes. *Lipids*, 52(5), pp.433–441. Available at: http://www.ncbi.nlm.nih.gov/pubmed/28161835 [Accessed March 28, 2018].

Callahan, H.L. et al., 1997. An axenic amastigote system for drug screening. *Antimicrobial agents and chemotherapy*, 41(4), pp.818–22. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9087496 [Accessed March 26, 2018].

Carrió, J. et al., 2000. Leishmania infantum: Stage-Specific Activity of Pentavalent Antimony Related with the Assay Conditions. *Experimental Parasitology*, 95, pp.209–214. Available at: http://www.idealibrary.com [Accessed March 9, 2018].

CDC, 2011. CDC - Neglected Tropical Diseases - Diseases. Available at: https://www.cdc.gov/globalhealth/ntd/diseases/index.html [Accessed December 6, 2017].

Croft, S.L., 1986. In vitro Screens in the Experimental Chemotherapy of Leishmaniasis and Trypanosomiasis. *64 Parasitology Today*, 2(3). Available at: https://ac-els-cdn-com.recursos.biblioteca.upc.edu/0169475886901572/1-s2.0-0169475886901572-main.pdf?_tid=3032cabe-e24b-4785-a21c-b6d5f49be9c5&acdnat=1520633242_30cf30f9b3878e9c8bef3b3c36a97023 [Accessed March 9, 2018].

Desjeux, P., 1996. Leishmaniasis: Public health aspects and control. *Clinics in Dermatology*, 14(5), pp.417–423. Available at: https://www.sciencedirect.com/science/article/pii/0738081X96000570 [Accessed January 6, 2018].

Ephros, M., Waldman, E. & Zilberstein, D., 1997. Pentostam induces resistance to antimony and the preservative chlorocresol in Leishmania donovani promastigotes and axenically grown amastigotes. *Antimicrobial agents and chemotherapy*, 41(5), pp.1064–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9145870 [Accessed March 26, 2018].

Escobar, P. et al., 2002. Sensitivities of Leishmania species to hexadecylphosphocholine (miltefosine), ET-18-OCH 3 (edelfosine) and amphotericin B. *Acta Tropica*, 81, pp.151–157. Available at: www.parasitology-online.com [Accessed March 9, 2018].

Ferrer, J., Prats, C. & López, D., 2008. Individual-based modelling: an essential tool for microbiology. *Journal of biological physics*, 34(1–2), pp.19–37. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19669490 [Accessed December 5, 2017].

Ferrer Savall, J. et al., 2010. *Individual-based modeling of plasmodium falciparum erythrocyte infection in in vitro cultures*. Universitat Politècnica de Catalunya. Available at: http://www.tdx.cat/handle/10803/6597 [Accessed March 8, 2018].

Feynman, R.P. (Richard P. & Bosch, A., 1967. *El carácter de la ley física*, Tusquets Editores. Available at: http://cataleg.upc.edu/search~S1*cat?/afeynman/afeynman/1%2C4%2C37%2CB/frameset&FF=afeynman+richard+p+richard+phillips+++++1918+++++1988&2%2C%2C34/indexsort=- [Accessed December 12, 2017].

Gállego, J., 2001. *Manual de Parasitología Morfología y biología de los parásitos de interés sanitario* Publicacions i Edicions UB, ed., Available at: http://www.publicacions.ub.es/refs/indices/06738.pdf [Accessed April 4, 2018].

Galvão, V., Garcia, J. & Miranda, V., 2009. Modeling the Chagas' disease after stem cell transplantation. *Physica A*, 388, pp.1747–1754. Available at: https://ac-els-cdn-com.recursos.biblioteca.upc.edu/S0378437109000387/1-s2.0-S0378437109000387-main.pdf?_tid=0edcbe21-acbc-485c-a75f-54892f291e8f&acdnat=1520635007_39c117cb21fa63fe3c57e0cf705bc83c [Accessed March 9, 2018].

Ginovart, M. et al., 2011. Modelos basados en el individuo y la plataforma NetLogo. *UNIÓN. Revista Iberoamericana de Educación Matemática*, 27, pp.131–150.

Goad, L.J., Holz, G.G. & Beach, D.H., 1984. Sterols of Leishmania species, implications for biosynthesis. *Molecular and Biochemical Parasitology*, 10(2), pp.161–170. Available at: https://www.sciencedirect.com/science/article/pii/0166685184900045 [Accessed March 28, 2018].

Gomes Neves, R. & Duarte Teodoro, V., 2010. Enhancing Science and Mathematics Education with Computational Modelling. *Journal of Mathematical Modelling and Application*, 1(2), pp.2–15. Available at: https://www.researchgate.net/profile/Rui_Neves2/publication/225041370_Enhancing_Science_and_Mathematics_Education_with_Computational_Modelling/links/0046352828d840e300000000/Enhancing-Science-and-Mathematics-Education-with-Computational-Modelling.pdf [Accessed December 12, 2017].

Górriz, M. et al., 2018. Leishmaniasis Parasite Segmentation and Classification using Deep Learning. In *X Conference on Articulated Motion and Deformable Objects*. (Accepted paper)

Grimm, V. et al., 2006. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 198(1–2), pp.115–126. Available at: https://ac-els-cdn-com.recursos.biblioteca.upc.edu/S0304380006002043/1-s2.0-S0304380006002043-main.pdf?_tid=8b8b3c9c-5ac2-4bd3-803c-53ab69c5dfdf&acdnat=1522256337_073284018bbe2710fdf25a01a7c9eb68 [Accessed March 28, 2018].

Grimm, V., 1999. Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future? *Ecological Modelling*, 115(2-3), pp.129–148. Available at: https://ac-els-cdn-com.recursos.biblioteca.upc.edu/S0304380098001884/1-s2.0-S0304380098001884-main.pdf?_tid=df54925e-eb1f-11e7-a993-00000aab0f01&acdnat=1514390958_abba2bde7f0c8153101ba1ca04de2441 [Accessed December 27, 2017].

Grimm, V. & Railsback, S.F., 2005. *Individual-based modeling and ecology*, Princeton and Oxford : Princeton University Press. Available at: http://cataleg.upc.edu/record=b1306430~S1*cat [Accessed December 11, 2017].

Hotez, P.J. et al., 2004. Combating Tropical Infectious Diseases: Report of the Disease Control Priorities in Developing Countries Project. *Clinical Infectious Diseases*, 38(6), pp.871–878. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14999633 [Accessed January 6, 2018].

Iman, R.L., Helton, J.C. & Campbell, J.E., 1981. An Approach to Sensitivity Analysis of Computer Models: Part I-Introduction, Input Variable Selection and Preliminary Variable Assessment. *Journal of Quality Technology*, 13(3), pp.174–183.

Kaiser, H., 1979. The dynamics of populations as a result of the properties of individual animals. *Fortschritte der Zoologie*, 25, pp. 53-74.

Mackey, T.K. et al., 2014. Emerging and reemerging neglected tropical diseases: a review of key characteristics, risk factors, and the policy and innovation environment. *Clinical microbiology reviews*, 27(4), pp.949–79. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25278579 [Accessed May 6, 2018].

Mckay, M.D., Beckman, R.J. & Conover, W.J., 2000. American Society for Quality A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code Output From a Computer Code. *Source: Technometrics*, 42(1), pp.55–61. Available at: http://www.jstor.org/stable/1271432 [Accessed April 11, 2018].

Mishra, J., Saxena, A. & Singh, S., 2007. Chemotherapy of Leishmaniasis: Past, Present and Future. *Current Medicinal Chemistry*, 14(10), pp.1153–1169. Available at: http://www.eurekaselect.com/openurl/content.php?genre=article&issn=0929-8673&volume=14&issue=10&spage=1153.

Moffat, J.G. et al., 2017. Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nature Reviews Drug Discovery*, 16(8), pp.531–543. Available at: http://www.nature.com/doifinder/10.1038/nrd.2017.111 [Accessed March 16, 2018].

Moreira, E.S., Anacleto, C. & Petrillo-Peixoto, M.L., 1998. Effect of glucantime on field and patient isolates of New World Leishmania: use of growth parameters of promastigotes to assess antimony susceptibility. *Parasitology research*, 84(9), pp.720–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9766900 [Accessed March 28, 2018].

Nagle, A.S. et al., 2014. Recent Developments in Drug Discovery for Leishmaniasis and Human African Trypanosomiasis. *Chemical Reviews*, (114), pp.11305–11347. Available at: https://pubs.acs.org/doi/pdf/10.1021/cr500365f [Accessed March 9, 2018].

Nelson, D.J. et al., 1979. Allopurinol ribonucleoside as an antileishmanial agent. Biological effects, metabolism, and enzymatic phosphorylation. *The Journal of biological chemistry*, 254(22), pp.11544–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/500658 [Accessed March 28, 2018].

Nühs, A. et al., 2015. Development and Validation of a Novel Leishmania donovani Screening Cascade for High-Throughput Screening Using a Novel Axenic Assay with High Predictivity of Leishmanicidal Intracellular Activity G. F. Späth, ed. *PLOS Neglected Tropical Diseases*, 9(9), p.e0004094. Available at: http://dx.plos.org/10.1371/journal.pntd.0004094 [Accessed March 8, 2018].

Ortiz, D. et al., 2017. Discovery of novel, orally bioavailable, antileishmanial compounds using phenotypic screening I. C. Almeida, ed. *PLOS Neglected Tropical Diseases*, 11(12). Available at: http://dx.plos.org/10.1371/journal.pntd.0006157 [Accessed March 9, 2018].

Peck, S.L., Simulation as experiment: a philosophical reassessment for biological modeling. Available at: https://ac-els-cdn-com.recursos.biblioteca.upc.edu/S0169534704002162/1-s2.0-S0169534704002162-main.pdf?_tid=5f32a0dc-eb21-11e7-a0ca-00000aab0f6c&acdnat=1514391602_b25d2029a8654ab1ba9b422838410806 [Accessed December 27, 2017].

Pedros Barnils, N., 2015. Analysis and modelling of Trypanosomatidae family in vitro and ex vivo cultures. Chagas disease and leishmaniasis. Available at: http://upcommons.upc.edu/handle/2117/77057 [Accessed December 5, 2017].

Peixoto, M.P. & Beverley, S.M., 1987. In vitro activity of sulfonamides and sulfones against Leishmania major promastigotes. *Antimicrobial agents and chemotherapy*, 31(10), pp.1575–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3435106 [Accessed March 28, 2018].

Railsback, S.F., 2001. Concepts from complex adaptive systems as a framework for individual-based modelling. *Ecological Modelling*, 139, pp.47–62. Available at: www.elsevier.com/locate/ecolmodel [Accessed December 27, 2017].

Rakeshsharma1 & Raviranjan2, R., ' Unculturable' bacterialdiversity: Anuntappedresource. Available at: https://www.researchgate.net/profile/RAVI_RANJAN/publication/216875765_Unculturable%27_bacterial_diversity_An_untapped_resource/links/0c97e75faace26b388710e47/Unculturable-bacterial-diversity-An-untapped-resource.pdf [Accessed December 11, 2017].

Rangel, H. et al., 1996. Naturally azole-resistant Leishmania braziliensis promastigotes are rendered susceptible in the presence of terbinafine: comparative study with azole-susceptible Leishmania mexicana promastigotes. *Antimicrobial agents and chemotherapy*, 40(12), pp.2785–91. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9124841 [Accessed March 28, 2018].

Raventós Roca, B., 2016. Desenvolupament de mètodes experimentals de cultius d ' amastigots de Leishmania spp . per millorar la quantificació dels resultats experimentals. Bachelor thesis at Universitat Politècnica de Catalunya (UPC). Available at https://upcommons.upc.edu/handle/2117/110645

Raventós Roca, B., 2017. Modelització matemàtica del Leishmania infantum Modelització matemàtica del comportament in vitro de Leishmania infantum. Master thesis at Universitat de Barcelona (UB).

Roberts, W.L., McMurray, W.J. & Rainey, P.M., 1998. Characterization of the antimonial antileishmanial agent meglumine antimonate (glucantime). *Antimicrobial agents and chemotherapy*, 42(5), pp.1076–82. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9593130 [Accessed March 28, 2018].

De Rycker, M. et al., 2013. Comparison of a high-throughput high-content intracellular Leishmania donovani assay with an axenic amastigote assay. *Antimicrobial agents and chemotherapy*, 57(7), pp.2913–22. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23571538 [Accessed March 14, 2018].

Samant, M., Kumar, A. & Pandey, S.C., 2018. Slow pace of antileishmanial drug development. *Parasitology Open*, 4(e4), pp.1–11. Available at: https://doi.org/10.1017/pao.2018.1 [Accessed May 27, 2018].

Sereno, D., Holzmuller, P. & Lemesre, J.L., 2000. Efficacy of second line drugs on antimonyl-resistant amastigotes of Leishmania infantum. *Acta Tropica*, 74, pp.25–31. Available at: www.elsevier.com/locate/actatropica [Accessed March 9, 2018].

Sharlow, E.R. et al., 2009. Identification of Potent Chemotypes Targeting Leishmania major Using a High-Throughput, Low-Stringency, Computationally Enhanced, Small Molecule Screen E. Ghedin, ed. *PLoS Neglected Tropical Diseases*, 3(11), p.e540. Available at: http://dx.plos.org/10.1371/journal.pntd.0000540 [Accessed March 16, 2018].

Siqueira-Neto, J.L. et al., 2010. Antileishmanial High-Throughput Drug Screening Reveals Drug Candidates with New Scaffolds. *PLoS Negl Trop Dis*, 4(5). Available at: https://www.dndi.org/wp-content/uploads/2010/05/Neto_May2010_Leish-promastigote-HTS.pdf [Accessed March 8, 2018].

Swinney, D.C., 2013. Phenotypic vs. Target-based drug discovery for first-in-class medicines. *Clinical Pharmacology and Therapeutics*, 93(4), pp.299–301.

WHO, 2017. WHO | World Health Organization. *WHO*. Available at: http://www.who.int/neglected_diseases/diseases/en/ [Accessed December 6, 2017].

Yamey, G. & Torreele, E., 2002. The world's most neglected diseases. *BMJ (Clinical research ed.)*, 325(7357), pp.176–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12142292 [Accessed January 6, 2018].

Yau, W.-L. et al., 2010. Cyclosporin A Treatment of Leishmania donovani Reveals Stage-Specific Functions of Cyclophilins in Parasite Proliferation and Viability B. L. Kelly, ed. *PLoS Neglected Tropical Diseases*, 4(6), p.e729. Available at: http://dx.plos.org/10.1371/journal.pntd.0000729 [Accessed May 27, 2018].

Annex A. ODD protocol for the description of the model.

Annex A. ODD protocol for the description of the model

The standard for description and communication of individual based model, the ODD protocol, described in Grimm et al. 2006 is used to describe this model. ODD stands for *Overview*, *Design concepts* and *Details* which will be described in detail below.

1. Overview

1.1. Purpose

The aim of the model is to study and describe the *in vitro* infection of macrophages of the RAW 264.7 cell line by *Leishmania infantum*. Adjusting the model to experimental data should help shed light on unknown mechanisms of action and on parameters that have not been measured or are difficult to determine experimentally. By readjusting the model to experimental data acquired from drug screening tests, the mechanisms of action of drugs could be inferred from their effects on the values of the model's parameters.

1.2. State variables and scales

The model considers two types of individuals or agents: macrophages and parasites. Each have multiple possible states. Macrophages can either be infected or not infected (healthy) and parasites can either be extracellular, adhered or intracellular. Parasites have the ability to adhere to macrophages and infect them. The state variables describing the elementary properties of these individuals are found in Table 1.

L. infantum is considered in promastigote or amastigote form depending on the stage of infection. When inoculated, all parasites are in promastigote form. Once infection occurs, those parasites that have successfully infected macrophage transform to the amastigote form. The washing of the culture ('3. Materials and Methods') is modelled by removing promastigotes at 24 h. The removal is not complete; therefore, it is modelled in a way that most promastigotes are removed, but not with 100 % efficiency. The model does not consider difference in behaviour between promastigotes and amastigotes, but it does consider differences in the parameters that govern them.

Table 1. State variables of agents.

Individual	Name	Description
Macrophage	diamac	Diameter of macrophage
	edatmac	Age of macrophage
	imac	Identification index of macrophage
	estatmac	State of macrophage (healthy or infected)
	tinfect	Time the macrophage has been infected
	repromac	Indicates if the macrophage is able to reproduce (1: yes, 0: no)
	rlogisind	R parameter of the logistic growth function of macrophages
Amastigote	diam	Diameter of amastigotes
	edatam	Age of amastigotes
	estatam	State of amastigotes (free, adhered or intracellular)
	inmac	Index of the macrophage to which adhered and intracellular amastigotes are linked.

The space modelled is equivalent to the optic field of a microscope and consists of a grid of 400x400 spatial cells. The model runs for 72 time steps and each step represents an hour. This is the length of the experiments carried out and no experimental data was obtained beyond 72 hours. In order to attain successful cultures after 72 hours, the culture medium must be renewed. This entails a proliferation of cells that makes quantification of infection difficult. If the culture is left for more than 72 h without medium renewal, the final state is very poor.

1.3. Process overview and scheduling

Time is measured discreetly and at each time step several actions are executed consecutively. Both macrophages and amastigotes get older with each time step. Macrophages grow according to a logistic function and can reproduce given they have sufficient diameter and surrounding space. If the number of parasites inside infected macrophages is larger than the maximum, they burst. The information generated by the simulation is collected and stored at each time step.

Free parasites can move around the space at random. If they come into contact with a macrophage, they can adhere to its surface and infect it. Failing to find and infect a macrophage entails a probability of death, related to the time spent in the free state. Inside macrophages, amastigotes can reproduce once they have reached a given age and will do so until the maximum number of amastigotes inside a macrophage is achieved.

First, actions related to macrophages are executed. The order of actions is: ageing, growth, reproduction and bursting. Next, the actions related to *Leishmania* are executed. The order of actions depends on the state of the parasite, so this must first be evaluated. If the state is extracellular, those who have exceeded the maximum amount of time in this state die. For those that do not die, the order of actions is: movement, adhesion, infection and reproduction. For those already intracellular, the only action a part from ageing is reproduction. The sequence of these actions is detailed in the model flow chart in Annex 2.

A significant process of the model, which mirrors the experimental protocol, is the removing of all free *Leishmania* after 24 hours. In the experimental method, all those parasites (promastigotes) that have failed to adhere to or infect a macrophage are removed by washing the culture, leaving only macrophages and intracellular amastigotes. This is reflected in the model by, at time step 24, evaluating which parasites have an extracellular state and by making them die, independent of the time they have spent in that state. The experimental process is not 100 % efficient and this is also taken into account: the vast majority of promastigotes is removed, but some are left behind.

The model uses periodic border conditions, which are default in the NetLogo platform. For each agent that leaves the modelled space, an equal appears on the opposite side of the space. This minimizes the effects of finite space and adds realism.

2. Design concepts

The dynamics of the community are not preconceived within the model but emerge from the interactions between individuals. This concept, emergence, has been discussed previously. The interaction between individuals, infection, is direct. There is no individual adaptation.

Stochasticity is included in the model at every level and it is essential for obtaining the appropriate emergent behaviours (Ferrer Savall et al. 2010). Individuals cannot perceive their surroundings and make decisions; therefore, their movement is random. There are also

probabilities related to death, adhesion and infection. Finally, all parameters of the model are subjected to Gaussian noise[1] with a value of 0.25.

L. infantum is considered in promastigote or amastigote form depending on the stage of infection. The physiological differences are only taken into account in certain parameters that govern their behaviour (time of death, distance of movement), but the model does not consider different rules for the two forms. When inoculated, all parasites are in promastigote form. Once infection occurs, those parasites that have successfully infected macrophage transform to the amastigote form.

3. Details

3.1. Initialisation

The individuals are created according to their properties and are distributed at random and without superposition in the area of 400x400 cells. All macrophages are healthy and all parasites are extracellular (in promastigote form).

There are a series of parameters that characterise the individuals and the simulation and that can be modified by the modeler. These are summarised in Table 2. Most of these parameters have not been determined experimentally nor have they been found in bibliography. Instead, they have been adjusted by comparing simulation results to experimental results.

Table 2. Simulation parameters that can be modified in the interface of the simulator.

	Parameters	Description	Source	Units
Macrophages	nmacini	Initial number of macrophages	Experimental method	
	diamacmax	Maximum diameter of macrophages	Parametrisation	µm
	rlogis	Parameter R of the logistic function describing the growth of macrophages	Parametrisation	h^{-1}
	tdmac	Duplication time of macrophages	(Assanga 2013)	h
	direpromac	Minimum diameter for reproduction of macrophages	Parametrisation	µm
	repart	Proportion of parasites divided between macrophages during duplication	Parametrisation	
	maxinfect	Maximum number of parasites per infected macrophage	Parametrisation	
Leishmania	namini	Initial number of parasites	Experimental method	
	diammax	Maximum diameter of parasites	(Gállego 2001)	µm
	tmortam	Time a parasite is able to spend outside, without infecting a macrophage	Parametrisation	h
	distam	Distance of movement of parasites	Parametrisation	µm
	probenganxa	Probability of adhesion of a parasite to a macrophage	Parametrisation	
	probinfecta	Probability of an adhered parasite of infecting a macrophage	Parametrisation	
	tdam	Duplication time of parasites	Parametrisation	h

[1] Gaussian noise is statistical noise having a probability density function equal to that of the normal distribution, which is also known as the Gaussian distribution. In other words, the values that the noise can take on are Gaussian-distributed.

The general subroutine for initialisation is *to setup*. It clears all variables and sets all global values to zero. It also creates the macrophages and distributes them in the space. In the experimental design ('3. Materials and Methods'), macrophages are cultured for 24 h by themselves before the promastigote culture is inoculated. The initial number of macrophages in the model is 100 and the initial number of promastigotes, 250. This is because in 24 h the initial 25 macrophages have reproduced twice (their duplication time is 12 h).

3.2.Inputs

As described in '3. Materials and Methods', the promastigotes are inoculated after the macrophages have been cultured for 24 h. This can be modelled as an input in the form of the subroutine *to sembra*. This subroutine creates the parasites according to their properties. Initially, all parasites are extracellular and distributed randomly in the space.

3.3.Sub-models

Three kinds of submodels can be distinguished: those related to macrophages, those related to *Leishmania* and those related to the culture system as a whole.

For macrophages:

i) Ageing

Macrophages increment their age at each time step. Their age is measured by the variable *edatmac*, whose units are time steps, each time step equivalent to one hour.

ii) Growth

Given they have enough space, macrophages grow according to the logistic function (1).

$$\frac{dD}{dt} = r \cdot \left(1 - \frac{D}{D_{max}}\right) \cdot D \quad (1)$$

In which r is the maximum growth rate (*rlogis*), D is the diameter of the macrophage (*diamac*) and D_{max} is the maximum diameter possible (*diamacmax*).

iii) Reproduction

Given they have enough space, when macrophages reach an age equivalent to the time of duplication (*tdmac*) and they have sufficient diameter (*direpromac*), they duplicate by bipartition.

For *Leishmania*:

i) Ageing

The parasites increment their age at each time step. When they change state between extracellular, adhered and intracellular, their age is set to zero and the count begins anew.

ii) Death

Extracellular and adhered parasites (those that have been unsuccessful in infecting a macrophage) have a certain probability of death related to their maximum age. This maximum age is defined by the parameter *tmortam*. Adhered parasites have double the maximum age with respect to free ones.

Because the simulator mirrors the experimental method described in '3. Materials and Methods', the culture is washed at 24 h. At all other time steps, the age of free parasites (equivalent to the time they have spent in this state) is evaluated and, if the maximum age (*tmortam*) has been reached, those parasites die. At time step 24, however, all extracellular parasites, independent of their age, die.

iii) Movement

The extracellular parasites can move in a random direction at each time step. The distance is determined by *distam*. The distance of movement is reduced to half once 24 hours have passed.

iv) Adhesion

Extracellular parasites have a certain probability, *probenganxa*, of adhering to the surface of a macrophage when they encounter one. It is indifferent if the macrophage is infected or not.

v) Infection

Parasites adhered to a macrophage have a certain probability, *probinfecta*, of infecting it.

vi) Reproduction

When the age of the intracellular parasites reaches the value of the duplication time, *tdam*, they can duplicate by bipartition until the maximum number of parasites per macrophage is attained.

Culture system:

i) Macrophage lysis

When the maximum number of parasites per macrophage (*maxinfect*) is reached, the macrophage bursts, returning the parasites to the extracellular state.

Annex B. Model code implemented in NetLogo

Annex B. Model code implemented in NetLogo

The following contains the code implemented in NetLogo that constitutes the described IBM of a *Leishmania infantum* and cell line RAW 264.7 *in vitro* culture.

```
;;;; Model of a Leishmania infantum and cell line RAW 264.7 in vitro culture ;;;;


extensions [csv]

;; infectable macrophages

breed [macs mac]

;; properties: diameter, age, identifier, state (0,1,2), infection (= 0, not infected), time spent infected, reproduction index (0 no, 1 yes)

macs-own [diamac edatmac imac estatmac infect tinfect repromac rlogisind]


;; infectious amastigotes

breed [ams am]

;; diameter, age, state, identifier of the cell it is related to (= 0, not related to any cell), age if it is infecting, state (0: free, 1: adhered, 2: infecting), to which macrophage it is adhered, which macrophage it has infected

ams-own [diam edatam edatam1 estatam inmac edatami]


;; global variables

;; time, number of macrophages, number of infected macrophages, number of amastigotes, number of adhered amastigotes, number of infecting amastigotes, mean number of amastigotes per macrophage

globals [temps nmacs nmacsinf nmacs2 nams namsengx namsinf nmacsmorts namsmorts  meanams


;; auxiliary variables

x0 y0 x1 y1 x2 y2 diam1 compta inc1 mou imac1 para aux1 aux2 distam1


;; parameters contained in the file

;; line with data, maximum diameter for macs, r of logistic growth function of macs, minimum diameter for reproduction of macs, time ams can spend in extracellular state, distance of movement of ams, probability of ams of adhering to a mac, probability of infecting mac, reproduction time of ams, proportion of ams divided up in mac division, maximum ams per mac

lhs diamacmax rlogis direpromac tmortam distam probenganxa probinfecta tdam repart maxinfect

;; number of lines (parameter combinations) contained in the file

runnum N
```

```
;; number of repetitions for each parameter combination

rep rep_go


;; auxiliary variables for the calculation of means for the repetitions for each combination

aux_inf ; % of infection

aux_d1 ; 1 i 2 ams per mac

aux_d2 ; 3 i 4 ams per mac

aux_d3 ; 5 or more ams per mac

lectura ; identifier for the parameter combination


;; output variables at time-points (12, 24, 36, 48 and 72 h)

inf_12 d1_12 d2_12 d3_12 inf_24 d1_24 d2_24 d3_24 inf_36 d1_36 d2_36 d3_36 inf_48 d1_48 d2_48
d3_48 inf_72 d1_72 d2_72 d3_72

]


;; Subroutine for the reading of the CSV file containing parameter combinations

to open-file

  clear-all

  file-close-all ; Close any files open from last run

  file-open "LHS.csv"

  set runnum 0

  set rep 30

  set N 2

end

; end of "open file" ;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;


;; Main script: reading of the file, assinging of parameter values and simulation

to read-file

  set runnum runnum + 1

  if (lectura >= N) [file-close stop]

  set lhs (csv:from-row file-read-line ";")

  show lhs

  set lectura item 0 lhs
```

```
set diamacmax item 1 lhs
set rlogis item 2 lhs
set direpromac item 3 lhs
set tmortam item 4 lhs
set distam item 5 lhs
set probenganxa item 6 lhs
set probinfecta item 7 lhs
set tdam item 8 lhs
set repart item 9 lhs
set maxinfect item 10 lhs
set rep_go 0
set inf_12 0
set d1_12 0
set d2_12 0
set d3_12 0
set inf_24 0
set d1_24 0
set d2_24 0
set d3_24 0
set inf_36 0
set d1_36 0
set d2_36 0
set d3_36 0
set inf_48 0
set d1_48 0
set d2_48 0
set d3_48 0
set inf_72 0
set d1_72 0
set d2_72 0
set d3_72 0
```

```
while [rep_go < rep]
  [
setup         ;; Initialization of simulation
go            ;; Simulation
  ]
export-files
end
; end of main script ;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;; Subroutine for the initialization of the simulation
to setup
  clear-turtles ; it is important to clear all agents from the previous simulation
  reset-ticks
  ask patches [set pcolor one-of [119]]

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;; intial cells
;; all are circles
  set-default-shape macs "circle"
  set-default-shape ams "circle"
;; indexes and globals set to zero
  set nmacs 0
  set nams 0
  set compta 0
  set nmacsmorts 0
  set namsmorts 0
  set temps 0
  set meanams 0
  set nmacsinf 0
  set namsinf 0
  set aux_inf 0
  set aux_d1 0
  set aux_d2 0
  set aux_d3 0
```

```
 set rep_go rep_go + 1


;; initial macrophages

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;; space

;; The cells occupy their circular space, without overlapping. The counter (contador) indicates how many times the simulator tries to allocate a cell in the space before giving up. In this sense, the value of the counter conditions how many cells are placed in the space.

  set compta 0

  set aux1 1

  while [(aux1 < nmacini) and (compta < 100000)]

      ;; positions and diameter are set at random, diamacmax is in CSV and noise is at the interface

     [set x1 random-float max-pxcor

      set y1 random-float max-pycor

      set diam1 diamacmax * (random-normal 1 noise)

      ;; cells are created so that they do not overlap

      ifelse all? macs [distancexy x1 y1 > (diam1 + diamac) / 2]

       [create-macs 1

        [set xcor x1

        set ycor y1

        set heading random 360

        set diamac diam1

        set repromac repro

        set rlogisind rlogis

        set infect 0

        set color 117

        set size diamac

        set aux1 aux1 + 1

        set nmacs nmacs + 1

        set imac nmacs]

        ]

      ;; if space is not found, the counter increases

       [set compta compta + 1]

      ]
```

```
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;; initial amastigotes (sembra)
sembra
end
; end of setup ;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;


;; Subroutine "sembra", inoculation of amastigotes
to sembra
;; diameter of ams is fixed by diammax (interface) and initially all ams are in free state
;; ams are placed in the space using the same procedure as macs (not overlapping)
  set diam1 diammax
  set aux1 0
  set compta 0
  while [ (aux1 < namini) and (compta < 10000)]
  [ set x1 random-float max-pxcor
   set y1 random-float max-pycor
   ifelse all? macs [distancexy x1 y1 > (diamac) / 2]
    [create-ams 1 [
      set xcor x1
      set ycor y1
      set diam diam1
      set color 115
      set size diam1
      set inmac 0
      set estatam 0
      set edatam 0
      set edatam1 (- tretard)
      set aux1 aux1 + 1
      set nams nams + 1]]
    [set compta compta + 1]
  ]
end
; end of "sembra" ;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
```

```
;; Simulation

to go

while [temps < tmax]

[

set temps temps + dt

ask macs

[set imac1 imac

 set x0 xcor

 set y0 ycor

 set edatmac edatmac + dt

 if (distmac > 0)

 [

  set repromac repro

  set compta 0

  set diam1 diamac

 ]


;;;;;;;;;;;;;;;;; MACROPHAGES

 set x1 xcor

 set y1 ycor

 let diam2 diamac + rlogis * (1 - diamac / diamacmax) * dt

 if all? other macs [distancexy x1 y1 >= (diam2 + diamac) / 2] [set diamac diam2 set size diamac]

; reproduction, if age and size are sufficient (with noise)

  if (edatmac > tdmac * (random-normal 1 noise)) and (diamac > direpromac * (random-normal 1 0.15))
and (repromac = 1)

  [

; reproduction if enough space is available

    set compta 002

    let diamac0 diamac

    let diamac1 (diamac / 2 ^ 0.5) * random-normal 1 0.05

    let diamac2 (diamac ^ 2 - diamac1 ^ 2) ^ 0.5

    let sumrad (diamac1 + diamac2) / 2

    set x0 xcor
```

```
  set y0 ycor
  set aux1 imac
  let angle random 360
  set x1 x0 + sumrad * cos(angle)
  set y1 y0 + sumrad * sin(angle)
  while [any? other macs with [distancexy x1 y1 < (diamac1 + diamac) / 2] and compta < 360] [set angle
angle + 1
  set compta compta + 1 set x1 x0 + sumrad * cos(angle) set y1 y0 + sumrad * sin(angle)]
  if compta < 360
   [
    set edatmac 0
    set diamac diamac1
    set size diamac
    set nmacs nmacs + 1
    hatch 1
    [set xcor x1 set ycor y1
     set diamac diamac2
     set size diamac
     set imac nmacs
     set aux2 imac
     set edatmac 0
     set repromac repro
     set infect 0
     set color 117]
    ask ams with [inmac = aux1]
    [ifelse (random-float 1 <= repart) [set heading towardsxy x0 y0 fd (diamac0 - diamac1) / 2]
    [ set xcor xcor + ( x1 - x0) set ycor ycor + ( y1 - y0 ) set heading towardsxy x1 y1 fd (diamac0 -
diamac2) / 2 set inmac aux2]]
    ask macs with [imac = aux1] [ set infect count ams with [(inmac = aux1) and (estatam = 2)]]
    ask macs with [imac = aux2] [ set infect count ams with [(inmac = aux1) and (estatam = 2)]]
    ]]
 ]
```

```
;;;;;;;;;;;;;;;; AMASTIGOTES

;;;;;;;;;;;;;;;; Actions of amastigotes / State 0: free; State > 0: related to a macrophage


;; WASHING

if temps = 24

[ask ams with [estatam = 0][if random-float 1 <= 0.25 [set nams nams - 1 die]]]

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

   ask ams with [estatam <= 1][set edatam edatam + dt

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;; adhered ams (estatam = 1) live double: tmortam * (1 + estatam)

   if (edatam > tmortam * (1 + estatam) * (random-normal 1 noise)) [set namsmorts namsmorts + 1 set
nams nams - 1 die]]

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;; those that are infecting

   ask ams with [estatam = 2]

    [set edatam1 edatam1 + dt

     set aux1 inmac

     if edatam1 > tdam * (random-normal 1 noise)

     [ask macs with [imac = aux1] [set x1 xcor set y1 ycor set infect infect + 1]

     set edatam1 0

     hatch 1 [

;; growing towards the centre, more or less

      set heading (towardsxy x1 y1)

      let desv1 180 * (random-normal 0 noise)

      set heading heading + desv1

      fd (distam / 2)

      set diam diammax

      set color black

      set size diam

      set estatam 2

      set edatam 0

      set edatam1 0

      set nams nams + 1]

     ]

    ]
```

```
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;; free ams can adhere ;;;
   ask ams with [estatam = 0][
   ifelse temps < 24 [set distam1 distam] [set distam1 0.5 * distam]
   set x1  xcor +  0.5 * distam1 * (random-normal 0 noise);
   set y1  ycor +  0.5 * distam1 * (random-normal 0 noise)
   set xcor x1
   set ycor y1
   set inc1 0
   ask macs
    [if (distancexy x1 y1 < diamac / 2) and (inc1 = 0)
     [if random-float 1 <= probenganxa
       [set inc1 imac
        set namsengx namsengx + 1
       ]]]
   if inc1 > 0 [set color yellow set inmac inc1 set estatam 1]
   ]
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;; adhered ams can infect
    ask ams with [estatam = 1]
    [if random-float 1 <= probinfecta
      [set aux1 inmac
       set estatam 2
       set color black
       ask macs with [imac = aux1][set infect infect + 1]]]

; Macrophage BURSTING
ask macs with [infect > maxinfect * (random-normal 1 noise)]
 [let iaux1 imac
  ask ams with [inmac = iaux1]
   [ set estatam 0
     set color 115
     set inmac 0
     set edatam 0
     set edatam1 0
```

```
  ]
  set nmacs nmacs - 1
 die
 ]
set nmacsinf count macs with [infect > 0]
set namsinf sum [infect] of macs
ifelse nmacsinf > 0 [
 set meanams namsinf / nmacsinf ]
[set meanams 0]

; Calculation of output data
set aux_inf ((count macs with [infect > 0]) / count macs) * 100
set aux_d1 100 * ((count macs with [infect = 1])+(count macs with [infect = 2]))/ count macs
set aux_d2 100 * ((count macs with [infect = 3])+(count macs with [infect = 4]))/ count macs
set aux_d3 100 * (count macs with [infect >= 5]) / count macs

assign
tick
]
end
; end of Simulation ;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;

;; Subroutine to assign output data to output variables
to assign
if (temps = 12)
[ set inf_12 inf_12 + aux_inf
 set d1_12 d1_12 + aux_d1
 set d2_12 d2_12 + aux_d2
 set d3_12 d3_12 + aux_d3
]
```

```
if (temps = 24)
[ set inf_24 inf_24 + aux_inf
  set d1_24 d1_24 + aux_d1
  set d2_24 d2_24 + aux_d2
  set d3_24 d3_24 + aux_d3
]

if (temps = 36)
[ set inf_36 inf_36 + aux_inf
  set d1_36 d1_36 + aux_d1
  set d2_36 d2_36 + aux_d2
  set d3_36 d3_36 + aux_d3
]

if (temps = 48)
[ set inf_48 inf_48 + aux_inf
  set d1_48 d1_48 + aux_d1
  set d2_48 d2_48 + aux_d2
  set d3_48 d3_48 + aux_d3
]

if (temps = 72)
[ set inf_72 inf_72 + aux_inf
  set d1_72 d1_72 + aux_d1
  set d2_72 d2_72 + aux_d2
  set d3_72 d3_72 + aux_d3
]
end
; end of "assign" ;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
```

```
; Subroutine to write output data to a CSV file
to export-files
;; write the information to the file
let spacer ","
  set inf_12 inf_12 / rep
  set inf_24 inf_24 / rep
  set inf_36 inf_36 / rep
  set inf_48 inf_48 / rep
  set inf_72 inf_72 / rep
  set d1_12 d1_12 / rep
  set d1_24 d1_24 / rep
  set d1_36 d1_36 / rep
  set d1_48 d1_48 / rep
  set d1_72 d1_72 / rep
  set d2_12 d2_12 / rep
  set d2_24 d2_24 / rep
  set d2_36 d2_36 / rep
  set d2_48 d2_48 / rep
  set d2_72 d2_72 / rep
  set d3_12 d3_12 / rep
  set d3_24 d3_24 / rep
  set d3_36 d3_36 / rep
  set d3_48 d3_48 / rep
  set d3_72 d3_72 / rep
 file-open (word "RESULTS1.csv")
 file-print (list spacer lectura spacer inf_12 spacer inf_24 spacer inf_36 spacer inf_48 spacer inf_72
            spacer d1_12 spacer d1_24 spacer d1_36 spacer d1_48 spacer d1_72
            spacer d2_12 spacer d2_24 spacer d2_36 spacer d2_48 spacer d2_72
            spacer d3_12 spacer d3_24 spacer d3_36 spacer d3_48 spacer d3_72 spacer)
 file-close
 file-open "LHS.csv"
 end
; end of "export files" ;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
```

Annex C. MATLAB code and flowchart for the Preliminary Analysis

Annex C. MATLAB code and flowchart for data analysis

The following contains a commented code implemented in MATLAB for the analysis of the simulation output of the NetLogo model. It imports the results of the model, calculates the mathematical distance and returns it as an output value. In particular, this code is used for '4.1. Preliminary analysis'. Its output is a matrix with each of the five evaluated values of a parameter and its corresponding mathematical distance. The flowchart corresponding to this code can be found in Figure C-1.

```
%% Calculation of mathematical distance to assess the proximity
between experimental and simulated values

clear all
close all

% Number of simulations for each parameter set
Nrep = 30;
% Output data: %Inf, 1 i 2 am, 3 i 4 am, >= 5 am
NVariablesOut = 4;
% Times at which output data is evaluated: 12h, 24h, 36h, 48h,
60h, 72h
NTempsAnalisi = 6;
% Number of input parameters to be studied
NParametres = 1;


%% Importation of simulated data
% The values of the input parameters are imported from a file
where all combinations are in order without being repeated

% Import the data
[~, ~, raw] = xlsread('parametres.xlsx','Hoja1','A2:B2');
% Path to where the parameters are stored

% Create output variable
data = reshape([raw{:}],size(raw));

% Allocate imported array to column variable names
Run = data(:,1);
original = data(:,2);

% Clear temporary variables
clearvars data raw;

NCombinacionsParametres = length(Run);
Parametres=[original];

% All experimental data is imported, but only the times 12h,
24h, 36h, 48h,60h and 72h are chosen

% As many runs as total simulations run
nRun = Nrep*NCombinacionsParametres;
% Number of rows in the matrix is the total number of
simulations
Nfiles = nRun;
```

```
% Number of columns is the six chosen times for each output
variable
Ncolumnes = NVariablesOut*NTempsAnalisi;
% Matrix to store the results
Resultats_simulacio = zeros(Nfiles,Ncolumnes);
% Loop for writing the chosen data to the matrix

iRun = 1;

for i=1:nRun

    % Importation of data
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

    filename = (['RESULTATS_',num2str(iRun),'.csv']);
    delimiter = ',';
    startRow = 2;


    %% Read columns of data as strings:
    % For more information, see the TEXTSCAN documentation.
    formatSpec = '%s%s%s%s%s%s%s%s%s%s%[^\n\r]';

    %% Open the text file.
    fileID = fopen(filename,'r');

    %% Read columns of data according to format string.
    % This call is based on the structure of the file used to
generate this
    % code. If an error occurs for a different file, try
regenerating the code
    % from the Import Tool.
    dataArray = textscan(fileID, formatSpec, 'Delimiter',
delimiter, 'HeaderLines' ,startRow-1, 'ReturnOnError', false);

    %% Close the text file.
    fclose(fileID);

    %% Convert the contents of columns containing numeric
strings to numbers.
    % Replace non-numeric strings with NaN.
    raw = repmat({''},length(dataArray{1}),length(dataArray)-1);
    for col=1:length(dataArray)-1
        raw(1:length(dataArray{col}),col) = dataArray{col};
    end
    numericData = NaN(size(dataArray{1},1),size(dataArray,2));
    for col=[2,3,4,5,6,7,8,9]
    % Converts strings in the input cell array to numbers.
Replaced non-numeric
    % strings with NaN.
    rawData = dataArray{col};
    for row=1:size(rawData, 1);
        % Create a regular expression to detect and remove non-
numeric prefixes and
        % suffixes.
```

```
        regexstr = '(?<prefix>.*?)(?<numbers>([-
]*(\d+[\,]*)+[\.]{0,1}\d*[eEdD]{0,1}[-+]*\d*[i]{0,1})|([-
]*(\d+[\,]*)*[\.]{1,1}\d+[eEdD]{0,1}[-
+]*\d*[i]{0,1}))(?<suffix>.*)';
        try
            result = regexp(rawData{row}, regexstr, 'names');
            numbers = result.numbers;
            % Detected commas in non-thousand locations.
            invalidThousandsSeparator = false;
            if any(numbers==',');
                thousandsRegExp = '^\d+?(\,\d{3})*\.{0,1}\d*$';
                if isempty(regexp(numbers, thousandsRegExp,
'once'));
                    numbers = NaN;
                    invalidThousandsSeparator = true;
                end
            end
            % Convert numeric strings to numbers.
            if ~invalidThousandsSeparator;
                numbers = textscan(strrep(numbers, ',', ''),
'%f');
                numericData(row, col) = numbers{1};
                raw{row, col} = numbers{1};
            end
        catch me
        end
     end
    end


    %% Split data into numeric and cell columns.
    rawNumericColumns = raw(:, [2,3,4,5,6,7,8,9]);
    rawCellColumns = raw(:, [1,10]);


    %% Allocate imported array to column variable names

    Temps = cell2mat(rawNumericColumns(:, 1));
    INFECCIO = cell2mat(rawNumericColumns(:, 2));
    Unam = cell2mat(rawNumericColumns(:, 3));
    Dosam = cell2mat(rawNumericColumns(:, 4));
    Tresam = cell2mat(rawNumericColumns(:, 5));
    Quatream = cell2mat(rawNumericColumns(:, 6));
    Cincam = cell2mat(rawNumericColumns(:, 7));
    meanamsmacr = cell2mat(rawNumericColumns(:, 8));

    %% Clear temporary variables
    clearvars filename delimiter startRow formatSpec fileID
dataArray ans raw col numericData rawData row regexstr result
numbers invalidThousandsSeparator thousandsRegExp me
rawNumericColumns rawCellColumns;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
% Transformation of simulator output variables

% Macrophages with 1 or 2 amastigotes
Unidosam = Unam + Dosam;
% Macrophages with 3 or 4 amastigotes
Tresiquatream = Tresam + Quatream;
% Parasitary index
IP = INFECCIO.*meanamsmacr;

     Aux1 = 0;
     Aux2 = 0;
     Raux = zeros(NVariablesOut,NTempsAnalisi);
     for j=1:NTempsAnalisi
        Aux1 = Aux1 + 1;
        Aux2 = Aux2 + 12;
     % Time interval selection
     Raux(1,Aux1) = INFECCIO(Aux2);
     Raux(2,Aux1) = Unidosam(Aux2);
     Raux(3,Aux1) = Tresiquatream(Aux2);
     Raux(4,Aux1) = Cincam(Aux2);
     Raux(5,Aux1) = meanamsmacr(Aux2);
     Raux(6,Aux1) = IP(Aux2);
     end

     % Data is placed in a results matrix (Resultats_simulacio)
that contains all the data for the mentioned output variables at
the time intervals 12h, 24h, 36h, 48h, 60h and 72h.

     ColIni = 1;
     ColFin = ColIni + NTempsAnalisi - 1;
     Resultats_simulacio(iRun,ColIni:ColFin) = Raux(1,:);

     ColIni = ColFin + 1;
     ColFin = ColIni + NTempsAnalisi - 1;
     Resultats_simulacio(iRun,ColIni:ColFin) = Raux(2,:);

     ColIni = ColFin + 1;
     ColFin = ColIni + NTempsAnalisi - 1;
     Resultats_simulacio(iRun,ColIni:ColFin) = Raux(3,:);

     ColIni = ColFin + 1;
     ColFin = ColIni + NTempsAnalisi - 1;
     Resultats_simulacio(iRun,ColIni:ColFin) = Raux(4,:);

     ColIni = ColFin + 1;
     ColFin = ColIni + NTempsAnalisi - 1;
     Resultats_simulacio(iRun,ColIni:ColFin) = Raux(5,:);

     ColIni = ColFin + 1;
     ColFin = ColIni + NTempsAnalisi - 1;
     Resultats_simulacio(iRun,ColIni:ColFin) = Raux(6,:);

     iRun = iRun + 1;

end
```

```
% The time interval 60h is eliminated because no experimental
data was obtained for this time interval

% Only the output variable %Inf, 1 i 2 am, 3 i 4 am, >= 5 am are
evaluated, but the model gives us more variables which must be
removed.

Resultats_simulacio(:,25:36) = [];

Aux3=0;
Aux4=5;
Aux5=0;

for i=1:4
    Aux3=Aux4-Aux5;
    Resultats_simulacio(:,Aux3)=[];
    Aux4=Aux4+6;
    Aux5=Aux5+1;
end

%% Importation of experimental data

Resultats_Experimentals = xlsread('ResultatsExperimentals');
NTempsAnalisiExp = 5;       % 12h, 24h, 36h, 48h i 72h

%% Calculation of the mathematical distance between experimental
and simulated values

% Point to point calulation of the mathematical distance between
the experimental and simulated results

distancia = zeros(nRun,20);

Aux6 = 1;

for k=1:4 % Number of columns of the matrix
"Resultats_experimentals"

for j=1:NTempsAnalisiExp

for i=1:nRun

    distancia(i,Aux6) = sqrt((Resultats_Experimentals(j,k) -
Resultats_simulacio(i,Aux6))^2);

end

Aux6 = Aux6 + 1;
end

end
```

```
% Distance of infection: related to the output variable %Inf

d_infeccio = distancia(:,1:5);

% Distance of distribution: related to how the amastigotes are
distributed within the macrophages (1 and 2 ams per mac, 3 and 4
ambs per mac, more than 5 ams per mac)

d_distribucio = zeros(nRun,NTempsAnalisiExp);

Aux7 = 6;
Aux8 = 11;
Aux9 = 16;

for j=1:NTempsAnalisiExp

for i=1:nRun

    d_distribucio(i,j) = sqrt((distancia(i,Aux7))^2 +
(distancia(i,Aux8))^2 + (distancia(i,Aux9))^2);

end

Aux7 = Aux7 + 1;
Aux8 = Aux8 + 1;
Aux9 = Aux9 + 1;

end

% Calculation of the mean for the evaluated times

distancia_mitjana = zeros(nRun,2);

for i=1:nRun
    distancia_mitjana(i,1) = mean(d_infeccio(i,1:5));
    distancia_mitjana(i,2) = mean(d_distribucio(i,1:5));
end

% Calculation of the global distance

distancia_global = zeros(Nrep,NCombinacionsParametres);

Aux10 = zeros(nRun,1);
for i=1:nRun
    Aux10(i,1) = sqrt((distancia_mitjana(i,1))^2 +
(distancia_mitjana(i,2))^2);
end

Aux11 = 1;
for j=1:NCombinacionsParametres
for i=1:Nrep
        distancia_global(i,j) = Aux10(Aux11,1);
Aux11 = Aux11 + 1;
end
end
 % Calculation of the mean of the global distance
```

```
distancia_global_mitjana = zeros(1,NCombinacionsParametres);

for i=1:NCombinacionsParametres
    distancia_global_mitjana(1,i) =
mean(distancia_global(1:Nrep,i));
end

%% Result
% Calculated distance with the value of the evaluated parameter

Resultat = zeros(2,NCombinacionsParametres);

Resultat(1,:) = Parametres(:,1);
Resultat(2,:) = distancia_global_mitjana(1,:);

% Results are stored in an xlsx file
xlswrite('Resultat.xlsx', Resultat)

%% Graphical output

figure(1)
boxplot(distancia_global(:,1),'Labels',{'Original'}) %Labels to
be modified according to parameter values
%ylim([0 60])   % Scale of axes to be adjusted according to
values of distance
hold on
title('Distància global')
xlabel('Valors paràmetres originals')
ylabel('Distància')
hold off
```

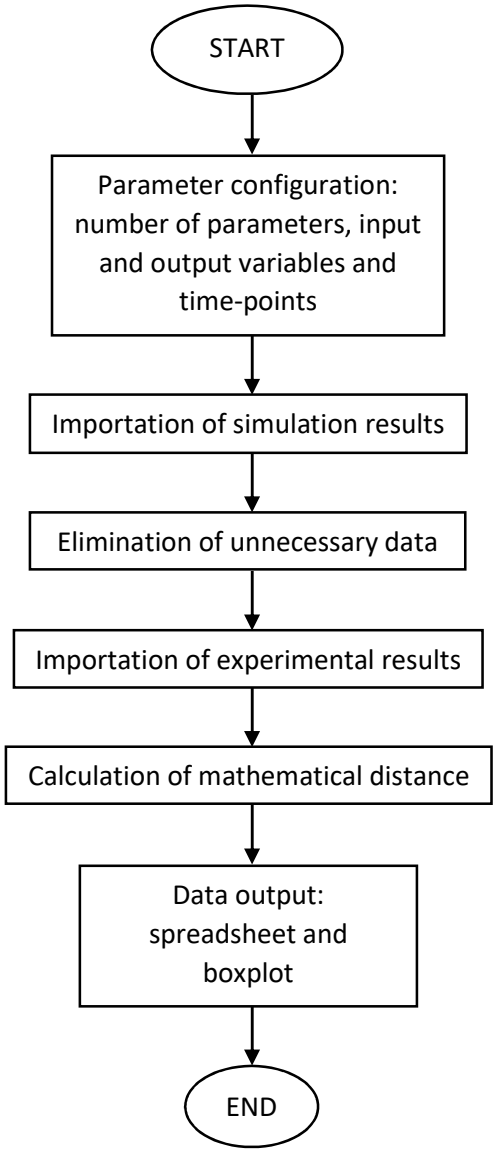


Figure C-1. Flowchart illustrating the MATLAB code used for the Preliminary Analysis of the IBM.

Annex D. MATLAB code and flowchart for Latin Hypercube Sampling

Annex D. MATLAB code and flowchart for Latin Hypercube Sampling

The following contains the code implemented in MATLAB for Latin Hypercube Sampling (LHS) of the parameter space of the model. In this case, the number of dimensions (parameters) is 10 and the number of samples, 1000. Figure D-1 shows the corresponding flowchart.

```
%% Parameter space sampling using the latin hypercube sample
method

%% Import parameter values
% Import lower and upper boundary values of parameters
[~, ~, raw] =
xlsread('Interval_parametres.xlsx','Intervals','B3:K4');

% Create output variable
data = reshape([raw{:}],size(raw));

% Allocate imported array to row variable names
lb = data(1,:);
ub = data(2,:);

%% Design of the LHS
p = 1000;         % Number of points
N = length(lb);  % Number of dimensions (parameters)
X = lhsdesign(p,N);
D = bsxfun(@plus,lb,bsxfun(@times,X,(ub-lb)));

%% Export to an Excel file
xlswrite('LHS.xlsx',D)
```

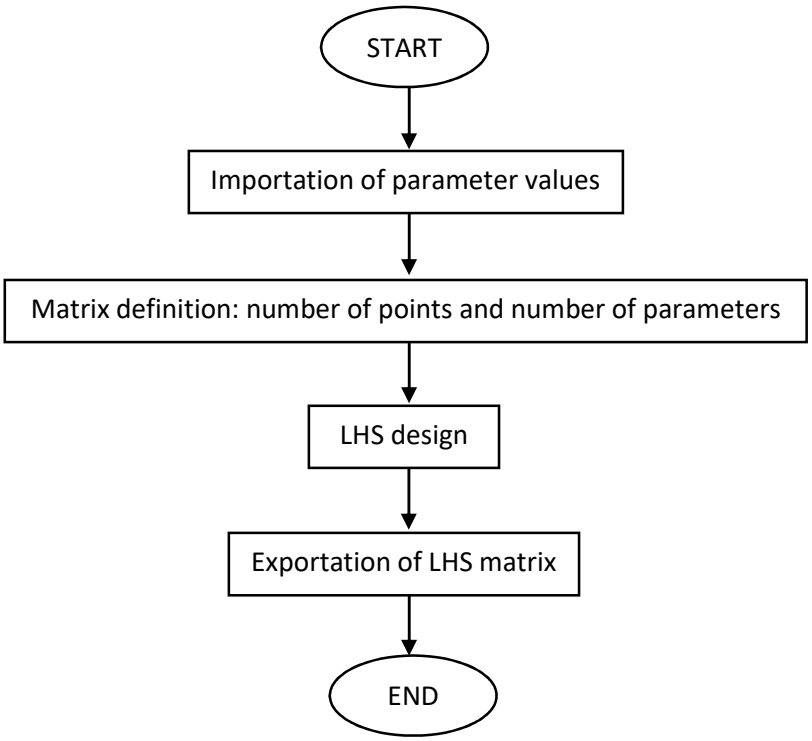


Figure D-1. Flowchart of the code used to sample the parameter space using the Latin Hypercube Sampling (LHS) technique.

Annex E. MATLAB code and flowchart for the analysis of the parameter space exploration

Annex E. MATLAB code and flowchart for the analysis of the parameter space exploration

The results of the parameter space exploration are analysed using the following code, implemented in MATLAB. The mathematical distance is calculated for each parameter combination. The output is an Excel file containing, in each line, the identifier of the parameter combination, the series of values of the parameters that make up the combination and the mathematical distance associated to implementing said combination in the simulator. The corresponding flowchart can be seen in Figure E-1.

```
%% Script for the analysis of the results from the automatized
simulations used for parametrisation with the LHS technique

clear all
close all

%% Parameters

NExpTimes = 5;                  % 12h, 24h, 36h, 48h i 72h
N = 1000;                       % Number of parameter combinations
NOutput = 4;                    % Number of output variables
                               (%inf,  1 and 2 ams, 3 and 4 ams,
                               more than 5 ams)


%% Importation of simulated data

 filename = (['RESULTS1_2605.csv']);
    delimiter = ',';
    startRow = 1;

%% Read columns of data as strings:
    % For more information, see the TEXTSCAN documentation.
    formatSpec =
'%s%s%s%s%s%s%s%s%s%s%s%s%s%s%s%s%s%s%s%s%s%s%s%s%s%[^\n\r]';

    %% Open the text file.
    fileID = fopen(filename,'r');

    %% Read columns of data according to format string.
    % This call is based on the structure of the file used to
generate this
    % code. If an error occurs for a different file, try
regenerating the code
    % from the Import Tool.
    dataArray = textscan(fileID, formatSpec, 'Delimiter',
delimiter,'ReturnOnError', false);

    %% Close the text file.
    fclose(fileID);
```

```
%% Convert the contents of columns containing numeric strings to
numbers.
    % Replace non-numeric strings with NaN.
    raw = repmat({''},length(dataArray{1}),length(dataArray)-1);
    for col=1:length(dataArray)-1
        raw(1:length(dataArray{col}),col) = dataArray{col};
    end
    numericData = NaN(size(dataArray{1},1),size(dataArray,2));
    for
col=[2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22]
    % Converts strings in the input cell array to numbers.
Replaced non-numeric
    % strings with NaN.
    rawData = dataArray{col};
    for row=1:size(rawData, 1);
        % Create a regular expression to detect and remove non-
numeric prefixes and
        % suffixes.
        regexstr = '(?<prefix>.*?)(?<numbers>([-
]*(\d+[\,]*)+[\.]{0,1}\d*[eEdD]{0,1}[-+]*\d*[i]{0,1})|([-
]*(\d+[\,]*)*[\.]{1,1}\d+[eEdD]{0,1}[-
+]*\d*[i]{0,1}))(?<suffix>.*)';
        try
            result = regexp(rawData{row}, regexstr, 'names');
            numbers = result.numbers;
            % Detected commas in non-thousand locations.
            invalidThousandsSeparator = false;
            if any(numbers==',');
                thousandsRegExp = '^\d+?(\,\d{3})*\.{0,1}\d*$';
                if isempty(regexp(numbers, thousandsRegExp,
'once'));
                    numbers = NaN;
                    invalidThousandsSeparator = true;
                end
            end
            % Convert numeric strings to numbers.
            if ~invalidThousandsSeparator;
                numbers = textscan(strrep(numbers, ',', ''),
'%f');
                numericData(row, col) = numbers{1};
                raw{row, col} = numbers{1};
            end
        catch me
        end
     end
    end

    %% Split data into numeric and cell columns.
    rawNumericColumns = raw(:,
[3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22]);
    rawCellColumns = raw(:, [1,10]);
    rawIndexColumns = raw(:, [2]);
```

```
%% Convert cells to numeric values
    Simulation_results = cell2mat(rawNumericColumns(:,:));
    Index = cell2mat(rawIndexColumns(:,:));


%% Importation of experimental data

Experimental_results = xlsread('Exp_results');


%% Calculation of mathematical distance
N_res = length(Experimental_results);
Distance = zeros(N,N_res);

for i=1:N_res
    for j=1:N
        Distance(j,i) = sqrt((Experimental_results(1,i) -
Simulation_results(j,i))^2);
    end
end

% Calculation of the mean value for distance for each outup
variable
Mean_distance = zeros(N,NOutput);

for i=1:N
    Aux1 = 1;
    Aux2 = NExpTimes;
    for j=1:NOutput
        Mean_distance(i,j) = mean(Distance(i,Aux1:Aux2));
        Aux1 = Aux1 + NExpTimes;
        Aux2 = Aux2 + NExpTimes;
    end
end

% Calculation of infection and distribution differences

d_inf = zeros(N,1);
d_dist = zeros(N,1);

d_inf(:,1) = Mean_distance(:,1);

for i=1:N
    d_dist(i,1) = sqrt((Mean_distance(i,2))^2 +
(Mean_distance(i,3))^2 + (Mean_distance(i,4))^2);
end

% Calculation of the global distance
Global_distance = zeros(N,1);

for i=1:N
    Global_distance(i,1) = sqrt((d_inf(i,1))^2 +
(d_dist(i,1))^2);
end
```

```
%% Import LHS matrix

Parameter_combinations = xlsread('Comparison');


%% Results
Results = zeros(N,12);
% 12 because 1 index, 10 parameter values and 1 distance columns

Results(:,1) = Index(:,1);
Results(:,2:11) = Parameter_combinations(:,2:11);
Results(:,12) = Global_distance(:,1);

xlswrite('Results',Results);
```

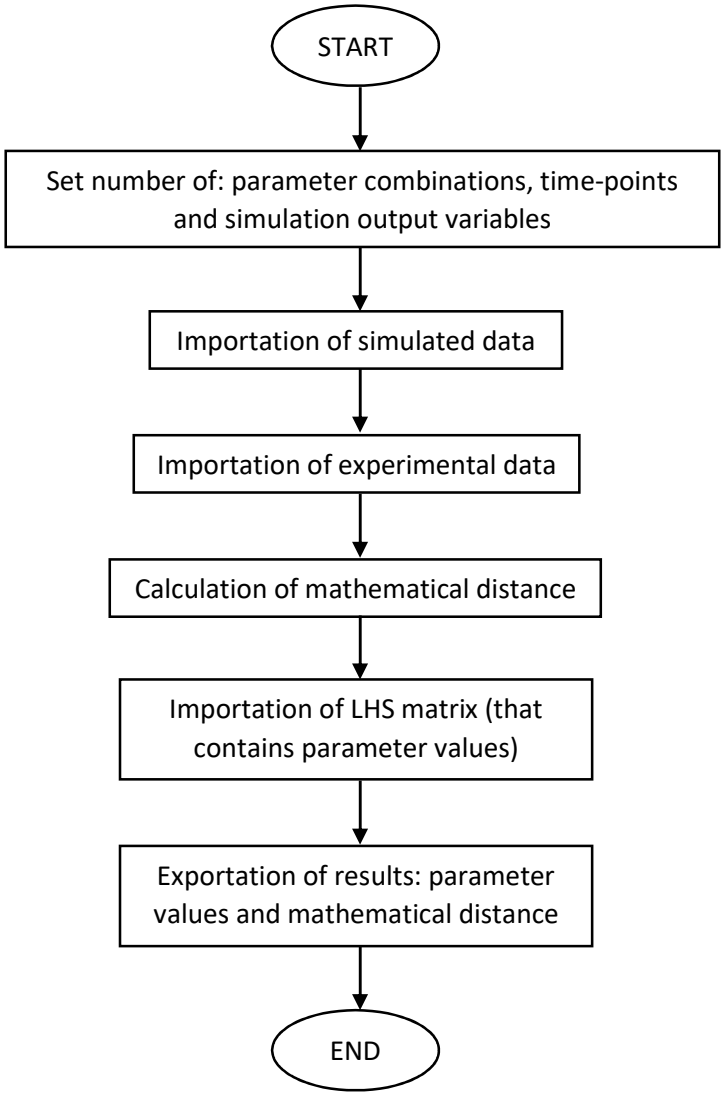


Figure E-1. Flowchart of the code used to analyse the simulation output of each parameter combination sample used in the exploration of the parameter space.